

TARTU ÜLIKOOL
MATEMAATIKA-INFORMAATIKA TEADUSKOND
Finants- ja Kindlustusmatemaatika

Anni Niidumaa

**Soome elektrienergia hinna modelleerimine ja
lühiajaline ennustamine**

Magistritöö

Juhendaja: dot. Raul Kangro

TARTU 2013

Olen Magistritöö kirjutanud iseseisvalt. Kõigile töös kasutatud teiste autorite töödele, põhimõttelistele seisukohtadele ning muudest allikaist pärinevatele andmetele on viidatud.

Autor: Anni Niidumaa

(allkiri)

.....

(kuupäev)

Sisukord

Sissejuhatus.....	5
1. Elektriturg	7
1.1 Mõned faktid elektrituru ja hinna kohta	7
1.1 Elektrienergia komponendid	9
1.2 Elektrienergia hinna kujunemine enne avatud elektriturgu ja avatud elektriturul.....	11
1.3 Elektrienergia hinna prognoosimine	12
2. Töös kasutatavad aegridade mudelid	14
3. Prognoosimine ainult minevikuhindasid kasutades	20
3.1 Soome elektrienergia originaalandmetele sobitatud lihtne mudel	20
3.1.1 Mudeli väljakirjutamine	26
3.2 Elektrienergia hinna logaritmitud andmete lihtne mudel	27
3.2.1 Logaritmitud andmete mudeli väljakirjutamine	31
3.3 Lihtsate mudelite võrdlus ja lühiajaline ennustamine	33
4. Prognoosimine täiendavaid regressoreid kasutades.....	36
4.1 Temperatuuri mõju Soome elektrienergia hindadele.....	37
4.1.1 Temperatuuri funktsiooni $Tt - \mu_2$ kasutamine regressor reana	38
4.1.2 Temperatuuri funktsiooni $Tt - \mu$ kasutamine regressor reana.....	43
4.1.3 Keerulise temperatuuri funktsiooni kasutamine regressor reana.....	47
4.2 Sademete mõju Soome elektrienergia hindadele.....	52
4.3 ARIMAX mudeli väljakirjutamine.....	61
4.4 Sobiva ARIMAX mudeliga lühiajaline ennustamine	63

Kokkuvõte.....	65
Summary	66
Kasutatud kirjandus	68
Lisa 1. Soome elektrienergia hinnad	70
LISA 2. Logaritmitud Soome elektrienergia hinnad	71
LISA 3. Keskmiste temperatuuri ja sademete tabel	72
LISA 4. Kasutatud skript R's	77

Sissejuhatus

Seoses üleminekuga avatud elektriturule on muutunud nii elektritootjate kui ka suurtarbijate jaoks äärmiselt oluliseks elektrienergia hindade prognoosimine. Teema olulisuse tõttu on seda küllalt palju uuritud ning on väljapakutud mitmesuguseid keerukaid mudeleid, mis üritavad arvestada kõiksuguseid mõjutegureid, nagu näiteks eri maade vahel olevate ühendustrasside läbilaskevõimed, elektrijaamade käivitamisega seotud kulud, seadusandlusest tulenevad mõjud, eri tüüpi elektrijaamade osakaalud ja palju muud. Samas keerulisi mudeleid on mõtet kasutada ainult siis, kui nende abil on saadavad prognoosid on oluliselt täpsemad, kui lihtsate mudelite abil on võimalik saada. Käesoleva magistritöö eesmärgiks ongi uurida, kui hästi on võimalik elektrihindade käitumist kirjeldada ning lühiajaliselt (kuni aasta ette) prognoosida mitmesuguste küllalt lihtsate mudelite abil. Kuna eesti kohta on seni liiga vähe andmeid, siis keskendume Soome elektrienergia hindade prognoosimisele.

Töö on jaotatud neljaks peatükiks. Esimeses peatükis tutvustatakse ja iseloomustatakse elektrituru olemust, hindade kujunemist nii fikseeritud turul kui ka avatud turul. Samuti on välja toodud peamised elektritootmise komponendid ja tootmisviisid.

Teises peatükis tutvustatakse esmalt töös kasutatavaid termineid, mida kasutatakse edasistes peatükkides.

Kolmandas peatükis vaadeldakse Soome elektrienergia hindade prognoosimist ainult minevikuhindade baasil, kasutades selleks ARIMA tüüpi mudeleid. Lisaks otse originaalandmetele sobitamisele uuritakse prognoosimist veel logaritmitud hindadele sobitatud ARIMA tüüpi mudelitega ning võrreldakse neid lähenemisi omavahel mitmesuguseid mõõdikuid kasutades.

Neljandas peatükis uuritakse võimalusi elektrienergia hinna prognooside parandamiseks juhul, kui võtta kasutusele täiendavad andmeid igakuiste keskmiste temperatuuride ja sademete kohta.

Modeleerimiseks kasutatakse antud töös tarkvarapaketti R, mis sisaldab vahendeid aegridade mudelite sobitamiseks, tulemuste graafiliseks esitamiseks ja leitud mudelite sobivuse kontrollimiseks. Tarkvarapakett R on laialdaselt kasutatav statistikas, finants- ja kindlustusvaldkonnas ning on kõikidele huvilistele tasuta allalaetav internetileheküljelt <http://www.r-project.org/>.

1. Elektriturg

Euroopa Liidus on toiminud avatud elektriturg juba alates 2007 aastast. Esimestena võtsid elektrituru kasutusele Rootsi, Soome ja Suurbritannia. Eesti ühinemisleping Euroopa Liiduga 2003 aastal tõi endaga kaasa mitmeid muudatusi ka siinsete elektri tootjate ja tarbijate jaoks. Leppega määrati kindlaks Eesti üleminek avatud elektriturule ja sellega seonduvad tingimused ja seadused. Eesti läks üle avatud elektriturule järk-järgult. Alates 2010 aastast võivad osta avatud elektriturult energiat suurtarbijad, st need kes vajavad elektrienergiat rohkem kui 2 GWh ning kes moodustavad vabatarbijatena kogu siseriikliku tarbimisest 35%. Tavatarbijatele, st väiketarbijatele avanes elektriturg 1. jaanuar 2013 [1].

Konkurentsiameti hinnangul on Eestis elektriturg äärmiselt kontsentreeritud ja orienteeritud ühele liigile, st ligikaudu 94% elektrienergiat toodetakse põlevkivist. Seda kontrollib sisuliselt suurim elektrienergiat tootev ettevõtte – Eesti Energia, kellele kuulub Eesti Põlevkivitööstus AS [10].

1.1 Mõned faktid elektrituru ja hinna kohta

Viimastel aastakümnetel on mitmed erinevad riigid tegelenud jõudsalt elektrienergia hinna alandamise ja efektiivsuse parandamisega, kusjuures Euroopa direktiivid annavad ette mitmed erinevad piirangud. Direktiivide eesmärgiks on edendada elektrituru liberaliseerimist. Viimane protsess on aeglane seoses erinevate seaduste kooskõlastamisega ning liberaliseerimine on tugevalt seostud majanduskasvuga. Mida kiiremini kasvab majandus seda jõudsamalt edendatakse ka turgusid. Elektriturg seab ette kauplemisega seotuid piiranguid – päevased maksimum kauplemiskogused,

kauplemise minimaalne hind. Kehtestatud on kõrged fikseeritud kauplemistasud neile, kes soovivad opereerida elektriturgudel [14].

Elektrituru avanemise protsess on mitmetes riikides juba ammu läbi, samas teistes riikides algus järgus. Eestis küll toimib avatud elektriturg, kuid siiski esineb probleeme, millele otsitakse jätkuvalt lahendusi. Sellest tulenevalt võime end selles valdkonnas lugeda teise arengufaasi riikide hulka.

Elektrienergia hinna käitumist on uuritud palju, kuid tänaseni ei ole veel leitud üldtunnustatud parimat mudelit, mis kajastaks võimalikult täpselt elektrihinna käitumist ning arvestaks adekvaatselt erinevaid tegureid (päevasisesed ja nädalapäevadega seotud efektid, temperatuur, sademed, alajaamade sisse ja väljalülitamised jne). Tänu sellele on elektrienergia hinna ennustamine suhteliselt ebatäpne.

Elektrienergia on põhimõtteliselt kaup, mille hinda iseloomustab kõrge varieeruvus. See tuleneb peamiselt faktist, et elektrienergiat ei saa hoiustada või ladustada, kuna see on enamasti kulukas ja majanduslikult mitte jätkusuutlik tegevus. Ainus mida saab vaadelda jätkusuutliku meetodina, on veereservide loomine, mida saab vajadusel kasutada elektrienergia tootmiseks. Muidugi veehoidlate rajamisega on seotud keskkonnaga seotud piirangud – pinnavorm, põhjavee asukoht ja asulate paiknevus [14].

Teine oluline asjaolu, mis mõjutab elektrienergia hinda on see, et ülekandevõrk pole kunagi täiuslik. Hinna varieeruvus eri piirkondades tuleneb ülekande, hoolduse ja seadmestiku maksumusest. Võimalikud ülekoormused ja potentsiaalsed rikked või tehnilised tõrked – need võivad äärmisel juhul viia süsteemi elektrikatkestusteni. Selliseid probleeme ei tohi kahe silma vahele jätta [13].

Elektriturul eristatakse kolme erinevat tüüpi perioodilisust. Esimene nendest on seotud suurema kutsliku valguse ja soojuse tarbimisega talvekuudel ja suurema õhukonditsioneeride kasutamisega suvekuudel. Teist tüüpi perioodilisus on nädalane perioodilisus, mis on põhjustatud argipäevase ja nädalavahetuse energia tarbimise erinevusest. Lõpuks päevasisene perioodilisus, mis tuleneb päeva ja öö erinevusest

ning ka päeva eristaadiumitest [14]. Arvesse tuleks võtta erinevaid relevantseid lisategureid, mille hulka kuuluvad näiteks ekstreemsed temperatuurid, keskkonna kataastroofid (üleujutused, maavärinad, vulkaanipursked, maanihked, tulekahjud jne), teatud tüüpi avalikud üritused ja protestid ning tehnilised probleemid (generaatorite rikked, elektriliinide rikked).

1.1 Elektrienergia komponendid

Elektrienergiat saab toota mitmest erinevast energiaallikast, millest osad on kallimad ja teised soodsamad, mõnest on lihtsam elektrienergiat kätte saada ning energiakaod väiksemad kui teistel. Peamisteks elektrienergia allikateks on järgmised:

Hüdroenergia- hüdroenergia on kõige odavam elektrienergia tootmise moodus ning seda on suhteliselt suurel määral nii Põhjamaade regioonis kui ka Eestis. Suurel määral mõjutab hüdroenergia elektrihipi just kevadisel ajal, mil veetase tõuseb märgatavalt. Veehoidlaid peetakse majanduslikult väga efektiivseteks, kuna need teenivad kiirelt kulud tagasi [7].

Tuuleenergia- Teine odav energia allikaks on tuuleenergia, mida Eestis kahjuks püsivalt ei ole. Tuulegeneraatoreid on ehitatud rannikualadele, millest pooled pole lülitatud Eesti Energia elektrivõrku seoses ettemääratud tingimuste mittetäitmisega. Tuuleenergia teeb hüdroenergiast kallimaks just kulukas energia salvestamine ning tuule juhuslikkusest tingituna esineb perioode, kus energia tootmine on negatiivne, st tuulegeneraatorid tarbivad rohkem energiat, kui suudavad toota. Eriti külmadel aegadel, mil temperatuur on -25 või külmem, lülitavad generaatorid ise end välja. Soodsama tuule puhul tarbivad toodetud elektrienergiat enamasti lähikondlased, st need elanikud, kes elavad tuulepargi naabruses [2].

Põlevkivi – Põlevkivi kaevandamine ei ole otseselt kulukas ettevõtmine, kuna põlevkivi asub 10-70 meetri sügavusel ning kaevandatav kihi paksus on keskmiselt 2,7-2,9 meetrit ning väheneb Lõuna Eesti suunas [5]. Põlevkivist energia tootmise teeb kalliks

mõju keskkonnale. Peale süsinikdioksiidi heitmemaksul tuleb arvesse võtta asjaolu, et kaevandused reostavad sissehingatavat õhku ja põhjavett [4]. Sellele vaatamata on põlevkivi hetkeseisuga põhiliseks kohalikuks energiaallikaks (katab 90% - 95% Eesti energiavajadusest) ning praeguste andmete alusel toimub veel kuni 2021 aastani põhiline elektrienergia tootmine põlevkivist. Eestis kaevandatakse suuremal määral põlevkivi Ida- Virumaal ning tootmine toimub Narvas.

Puit – Puit on taastuv loodusressurss, mida kasutatakse suhteliselt vähe elektrienergia tootmiseks. Euroopa Liidu Nõukogu 23.aprilli 2003 direktiiv näeb ette taastuva energia kasutamist 25% ulatuses kogu toodetavast energiast, mis tähendab omakorda puidutööstuse nõudluse kasvu. Hinnanguliselt läheb vaja 30- 50% enam puitu kui praegu [2]. Majanduslikult võttes on puidust energia tootmine kallim kui põlevkivist [11], kuid keskkonnasõbralikum ning elavdab elu maapiirkondades [6].

Turvas – Turvast ei loeta taastuvate loodusvarade hulka, vaid aeglaselt taastuvate loodusvarade hulka. Turvast kasutatakse suuremas osas soojusenergia tootmiseks, kuna turba kaevandamiseks on ettemääratud kvoodid, millest peab kinni pidama ning energia tootmisel tuleb tasuda CO₂ heitmemaksu. Turba kasutamist energeetikas raskendab asjaolu, et turba lagunemisastmest sõltub tema energeetiline väärtus ning väikene mahukaal raskendab transportimist ja ka põletamist [9, lk 35-36]. Turba kasutamisega kaasnevad mitmed lisakulud, mis on peamiselt seotud keskkonnateguritega (taimestiku eemaldamine, kändude eemaldamine kaevandamise alalt, kuivenduskraavide rajamine, tekkinud vee kontrollimine ja vajadusel filtreerimine enne loodulikku veekogusse juhtimist). Turba kaevandamine on hooajaline tegevus (varakevadest peale lume ja maapinna sulamist kuni augustini) [9, lk 69-71], millest tingituna pole turvas hetkelises majandusolukorras jätkusuutlik energiatoomise allikas.

Tuumaenergia – Tuumaenergiat loetakse säästvaks energiaks, kuna tootmise käigus ei eraldu CO₂, kuid kaasneb oht radioaktiivse saaste kandumiseks keskkonda. Nagu kõik teised energiaallikad, saastab ka tuumaenergia keskkonda, nimelt tuumajäätmete ladustamisel sügavale maa- alustesse hoidlatesse – see kõik võib rikke puhul reostada põhjavett ning ka hingatavat õhku [19].

Tuumaenergiat saadakse tuumkütusest (uraanimaagist), mida peamiselt Euroopasse tuuakse sisse Austraaliast [18]. Tuumaenergiajaamade ehitamist peetakse kalliks seoses kõrgete nõuetega ning aeganõudvama ehitamisega kui teiste energiatootmisjaamade puhul [20].

Maagaas – Maagaas tuuakse sisse Venemaalt Kaliningradi oblastist ning Baltimaades on maagaasi osakaal kogu elektrienergia tootmisest 25%. Maagaasist elektri tootmise teeb kalliks transportimine ja sellega kaasnevad riskid ning hoiustamine Lätis Incukalnsi maa-aluses hoidlas.

1.2 Elektrienergia hinna kujunemine enne avatud elektriturgu ja avatud elektriturul

Elektrienergia hind reguleeritud turul oli määratud lepingutega, kus puudub arbitraaži võimalus, st puudus ühe hinna seadus. Elektriturul pakkuja sai tõsta ja langetada hinda vastavalt nõudlusele ning sellest tulenevalt olid vastavate graafikute joonköverad jäigad, st ühe joongraafiku muutus ei too kaasa teise joongraafiku muutumist [1].

Enne elektrituru avanemist arvutati elektrihipa kulupõhiliselt ning fikseeriti hind kolmeks järgneva aastaks, mis kooskõlastati Konkurentsiameti poolt. Tavatarbija jaoks tähendas see fikseeritud hinda.

Kulupõhiline arvestamine lähtub plaanilistest töödest ja investeeringute mahust. Kui vaadelda viimase kolme aasta elektrihipa, siis on see olnud suhteliselt kõrge. See oli tingitud suurtest ja mahukatest investeeringutes nagu seda on EstLin1 välja ostmine, EstLink 2 ehitamine, avariikeskuse rajamine Kiisal ning ühe 330 kW elektriliini ehitamisest. Kulude sisse on arvestatud veel plaanilised alajaamade rekonstrueerimised ja renoveerimised, liinide taastamised ja maaaluste kaablite asendamised uutega. Kõik see tehakse selleks, et igale tarbijale jõuaks igal ajal elekter, st et oleks tagatud elektrivarustuse kindlus. Nagu eelpool sai mainitud, oli varasemalt elektrihipa määratud kindlaks mingiks perioodiks. Elering ja Eesti Energia tegid

omapoolsed pakkumised elektrihinna kujunemise kohta Konkurentsiametile ning kui kõik nõuded ja tingimused olid täidetud, siis kinnitas Konkurentsiamet uued hinnad. Vastasel juhul tuli organisatsioonidel teha hinna suhtes järelandmisi ja seejärel uuesti esitada.

Elektrituru avanemise peamine eesmärk on konkurentsi tekitamine võimalikult mitmes elektrienergia tarnimise lülis. Turu avanemisega tekkis turuosalistele mitmeid uusi võimalusi tegutsemiseks, tekkis konkurents elektritootmises ja elektrikaubanduses, samal ajal jäävad võrkude infrastruktuuride ja süsteemiteenustega seotud funktsioonid ka edaspidi monopolseteks, sest mitme paralleelse elektriliini ehitamine on majanduslik nonsens [3].

Peamiseks elektrienergia hinda mõjutavaks teguriks avatud elektriturul ja elektrikörsil on piisavate tootmisvõimsuste ning elektriühenduste olemasolu, et tagada elektri liikumine nii siseriiklikult kui naaberriikidega [14]. Tavatarbijate jaoks suurimaks muudatuseks on vabadus valida pakkuja, kes pakub antud hetkel ja tingimustel parima hinnaga elektrienergiat.

Hind avatud turul kujuneb pakkumise ja nõudluste kokkulangemispunktis ning seda arvutatakse iga tunni järel, võttes kokku kõik pakkumised. Üldiselt võttes: mida suurem on nõudlus, seda kõrgem on elektrienergia hind. Nõudluse suurenemine määrab ära ka kallimate energeetiliste ressursside kasutamise energia tootmisel. Teisisõnu tarbimise kasvades suureneb ka tootmisemaht, mis toob endaga kaasa toodangu marginaalkulude kasvu. See omakorda kasvatab hüdroenergia, nafta, gaasi, tuumaenergia ja ka kivisöe kasutamist eksponentsiaalselt.

1.3 Elektrienergia hinna prognoosimine

Elektrienergia hinna uurimine ja sobiva mudeli leidmine on päevakorraline teema, millega tegelevad paljud oma ala spetsialistid. Varasematest avaldatud töödest üks osa kajastab elektrienergia hinna kujunemise majanduslikke aspekte ja teine osa

keskenduvad prognoosidele, mis tuginevad elektrienergia hindade empiirilistele analüüsidele. Samuti on tähelepanu pööratud metoodilistele aspektidele – peamiselt küll majanduslikele kuid siiski vaadeldakse ka statistilisi aspekte, kus hinnatakse mudeli tõhusust sõltuvalt hetkehinna ennustusvigadest [14]. Analüüs keskendub peamiselt Nord Pool Spot'i elektrituru energia hinna väärtustele, mis on ühtlasi Euroopa suurim tegutsev elektrienergia turg.

Elektrienergia hinna määramisel on kasutatud mitmeid erinevaid mudeleid, millest osad arvestavad lihtsalt elektrienergia hinna ajaloolisi väärtusi. Samas on mudeleid, mis võtavad arvesse mitmeid erinevaid välistegureid, nagu elektrijaamade sisse ja välja lülitamisega seotus lisakulud, võimalike elektrikatkestused, keskkonnategurid – temperatuuri kõikumisi, sademete hulk, tuule kiirus jne. Lisaks on uuritud heteroskedastilisi mudeleid, st mudeleid, mille jääkide dispersioonid ei ole sõltumatud ning hinna varieeruvus ajas on muutuv.

2. Töös kasutatavad aegridade mudelid

Enne kui hakkame leidma Soome elektrienergia hindadele sobivaid mudelid, defineerime mõisted, mida kasutame antud töös.

Definitsioon 1: Protsessi Z_t , $t \in \mathbb{Z}$, nimetatakse teist järku nõrgalt statsionaarseks protsessiks, kui $D(Z_T) = \gamma_0$ ning $cov(Z_t, Z_{t+p}) = \gamma_p$ iga t korral, kus γ_p ($p \in \mathbb{Z}$) on mingid t -st mitte sõltuvad reaalarvud.

Käesolevas töös tähendab mõiste „statsionaarne protsess“ teist järku nõrgalt statsionaarset protsessi.

Andmetele on tihti peale võimalik sobitada mitu erinevat mudelit, mis võivad näiteks erineda üksteisest parameetrite arvu poolest. Parameetrite arvu on võimalik lähtuda autokorrelatsioonidest ja või osautokorrelatsioonidest.

Definitsioon 2: Autokorrelatsiooniks nimetatakse statistikas lineaarseks sõltuvuseks kahe eri tunnuse vahel. Statsionaarse protsessi puhul autokorrelatsioon sõltub ainult ajamomentide vahest p ja defineeritakse

$$\rho_p = cor(Z_t, Z_{t+p}) = \frac{cov(Z_t, Z_{t+p})}{\sqrt{Var(Z_t)Var(Z_{t+p})}} = \frac{\gamma_p}{\gamma_0},$$

Kus Z_t on juhuslik protsess [15].

Definitsioon 3: Statsionaarse protsessi Z_t k -ndat järku osautokorrelatsiooni-kordajaks nimetatakse suuruste Z_t ja Z_{t-k} vahelist korrelatsiooni pärast suuruste $Z_{t-1}, \dots, Z_{t-(k-1)}$ mõju eemaldamist [16].

Eelnevas definitsioonis mõistetakse mõju eemaldamise all vastavate suuruste abil moodustatud parima lineaarse prognoosi mahalahutamist.

Tihti kasutatakse aegridade statistilises analüüsis ARMA mudeleid, mis koosnevad kahest osast, milleks on autoregresiivne osa ja liikuva keskmise osa.

Definitsioon 4: Eeldame, et juhuslikud suurused $A_t, t \in \mathbb{Z}$ on tsentreeritud sõltumatud sama jaotusega juhuslikud suurused. Olgu μ juhusliku protsessi Z_t väärtuste keskvärtus. Kasutame tähistust $\tilde{Z}_t = Z_t - \mu$. Defineerime järgmised lõpliku arvu kordajatega lineaarsete protsesside klassid:

- Järguga p autoregresiivseks protsessiks ehk $AR(p)$ protsessi nimetatakse protsessi kujul

$$\tilde{Z}_t = \sum_{i=1}^p \varphi_i \tilde{Z}_{t-i} + A_t,$$

kus parameetrid $\varphi_1, \dots, \varphi_p$ nimetatakse AR koefitsientideks ja liiget A_t nimetatakse häirituseks.

- Järguga q liikuva keskmisega protsessiks ehk $MA(q)$ protsessi nimetatakse protsessi kujul

$$\tilde{Z}_t = A_t - \sum_{i=1}^q \theta_i A_{t-i}$$

kus parameetreid $\theta_1, \dots, \theta_q$ nimetatakse MA koefitsientideks ja suuruseid A_t häiritusteks.

- $ARMA(p,q)$ protsessiks nimetatakse protsessi kujul

$$\tilde{Z}_t = \sum_{i=1}^p \varphi_i \tilde{Z}_{t-i} + A_t - \sum_{i=1}^q \theta_i A_{t-i}.$$

Sobivate mudelite otsimisel on kasuks teadmine, et $AR(p)$ protsessi korral on osautokorrelatsioonid võrdsed nulliga alates järgust $p+1$ ning $MA(q)$ protsessi korral on autokorrelatsioonid võrdsed nulliga alates järgust $q+1$.

Sageli on aegrea keskmine ajas muutuv ja seega ei vasta rida statsionaarsuse nõuetele. Samas võib aga muutude protsessi kirjeldamiseks sobida mõni AR , MA või $ARMA$ tüüpi mudel. Selliseid protsesse nimetatakse $ARIMA$ (autoregresiivne integreeritud liikuva keskmisega) protsessideks [16]. Diferentside kirjapanekus on kasulik defineerida tagasinihke operaator, mida tähistatakse

$$B(Z_t) = Z_{t-1}.$$

Tehniliselt tähendab seda, et nihkeoperaator B teisendab protsessi Z uueks protsessiks, mille t -s element on esialgse protsessi väärtus kohal $(t - 1)$.

Definitsioon 5: ARIMA(p,d,q) protsessiks nimetatakse juhusliku protsessi Z_t , mille d-ndat järku muudud (diferents) $W_t = (1 - B)^d Z_t$ on esitatav kujul

$$\tilde{W}_t = \sum_{i=1}^p \phi_i \tilde{W}_{t-i} + A_t - \sum_{i=1}^q \theta_i A_{t-i},$$

kus $\tilde{W}_t = W_t - EW_t$ ning juhuslikud suurused A_t on statsionaarsed, mittekorreleeritud ning sõltumatud suurustest $W_{t-i}, i = 1, 2, \dots$.

Definitsioon 6: Olgu meil statsionaarsed aegread Z_t ja $X_{it}, i = 1, 2, \dots, m$, siis mitmese lineaarse regressiooni mudeliks nimetatakse mudelit kujul

$$Z_t = \beta_0 + \sum_{i=1}^m \beta_i X_{it} + \varepsilon_t,$$

kus ε_t on sama jaotusega, sõltumatud ja tsentreeritud juhuslikud suurused. Kui vead ε_t vastavad mingile ARMA protsessile $\phi(B)\varepsilon_t = \theta(B)A_t$ ning häiritused A_t on sõltumatud protsessi X_t väärtustest, siis vastavat mudelit nimetatakse ARIMAX tüüpi mudeliks, mis on kujul

$$Z_t = \beta_0 + \sum_{i=1}^m \beta_i X_{it} + \varepsilon_t.$$

ARIMAX mudelite erijuhuks on ülekandefunktsiooni mudelid, kus eeldatakse, et Z_t ja X_t on statsionaarsed protsessid ning regressoriteks X_{it} on sama aegrea nihked.

Definitsioon 7: ARIMAX mudeliks nimetatakse mudelit kujul

$$Z_t = \beta_0 + \sum_{i=b}^{\infty} \beta_i X_{t-i} + \varepsilon_t,$$

kus

$$\beta(x) = \beta_0 + \sum_{i=b}^{\infty} \beta_i x^i$$

avaldub kujul

$$\beta(x) = \beta_0 + x^b * \frac{(v_0 + v_1 * x + \dots + v_s * x^s)}{(1 - \delta_1 * x + \delta_2 * x^2 + \dots + \delta_r * x^r)} = x^b \frac{\sum_{i=0}^s v_i x^i}{1 - \sum_{i=1}^r \delta_i x^i}.$$

Eelmist mudelit on võimalik kirjutada kujule

$$Z_t = \sum_{i=1}^r \delta_i Z_{t-i} + \sum_{i=0}^s v_i X_{t-b-i} + \varepsilon_t,$$

kus suurused ε_t vastavad ARMA protsessile kujul $\phi(B)\varepsilon_t = \delta(B)\theta(B)A_t$ ning B tähistab tagasinihke operaatorit ning suurused δ_i, ϑ_i on mingid arvud, $\delta(B)$ on vastav polünoom ning $\phi(B)$ ja $\theta(B)$ on polünoomid.

Definitsioon 6 ütleb, et häiritused A_t peavad olema sõltumatud protsessi X_t väärtustest, siis on ka suurused ε_t sõltumatud suurustest X_t . ARIMAX mudeli leidmine on kaheetapiline: esimeses osas sobitame mitmese regressiooni mudeli vigade leidmiseks, teises osa leiame vigade jaoks sobiva mudeli kuju ja sobitame andmetele leitud kujule vastava ARIMAX mudeli. Ülekandefunktsiooni mudeli korral eelneb eelmainitud tegevusele veel sobiva nihke b ning parameetrite s ja r leidmine.

Kui juhuslik protsess ei ole statsionaarne (näiteks keskvärtus on ajas muutuv), siis tuleb kas aegreast eemaldada trend või vastavat rida diferentsida. Samas diferentsimine vähendab sobitamiseks kasutatavat andmestikku ja liigselt diferentsitud reale sobitatud mudel on ebatäpsem, kui õige arv diferentsitud reale sobitatud mudel. Statsionaarsust saab kontrollida autokorrelatsioonide graafikult, kui vastava aegrea autokorrelatsioonide väärtused püsivad ühe lähedal, siis see viitab rea mitte statsionaarsusele. Sellest tulenevalt tuleks rakendada diferentsimist. Sesonset diferentsi on mõistlik kasutada, siis kui aegrea autokorrelatsioonid esinevad mingi sammuga, mis on ka vastavaks perioodiks.

Et aegrida mitte üleliia diferentsida, siis on mõistlik kasutada Phillips-Perron testi, seda on võimalik kasutada ka paketi R.

Phillips - Perron testi kasutatakse ARIMA tüüpi mudelite statsionaarsuse kindlaks tegemiseks, kus nullhüpoteesiks on vastava aegrea diferentsimine (st, et esialgne rida ei ole ka peale lineaarse trendi eemaldamist statsionaarne, kuid diferentsitud rida on statsionaarne). Alternatiivseks hüpoteesiks on see, et peale lineaarse trendi eemaldamist, on rida statsionaarne. Lihtsamalt öeldes, kasutades paketi R vastavat testi saame mingi p-väärtuse. Kui vastav väärtus on väiksem kui 0.05, siis uuritavat rida ei ole mõistlik diferentsida. Siinkohal peab meeles pidama, et test ei tööta sesoonse diferentsi leidmise vajaduse korral [12].

Sageli esineb olukordi, kus ühele ja samale andmestikule sobib mitu erinevat mudelit, mille parameetrite arv ei pruugi olla sama, siis alati jääb õhku küsimus: milline sobivatest mudelitest on parim. Sellisele küsimusele aitab leida vastuse Akaike informatsiooni kriteerium (lühemalt AIC), kuna AIC arvestab nii parameetrite arvu kui ka suurimat tõepära ning mis on leitav valemiga

$$AIC = 2k - 2 \ln(L),$$

kus k tähistab mudeli parameetrite arvu ning L tähistab aegreale sobitatud mudelile vastavat tõepärafunktsiooni väärtust. Kriteerium aitab välistada andmete ülesobitamist. Seega mida väiksem on Akaike informatsiooni kriteeriumi väärtus seda sobivam on mudel [8].

Sarnaselt AIC aitab mudeli sobivust kontrollida Bayesian Informatsiooni kriteerium (lühemalt BIC) kordaja, mis on arvutatav valemiga

$$BIC = -2 \ln(L) + 2k \ln(n),$$

kus L on vaatluste tõepära sobitatud mudeli korral, k tähistab parameetrite arvu ning n aegrea vaatluste arvu. BIC on väga tugevalt seotud AIC väärtusega ning BIC valem arvestab parameetrite arvu tugevamini kui AIC. Seega BIC väärtuse põhjal valitud mudelil ei ole parameetrite arv kunagi suurem kui AIC põhjal valitud mudelil. Meeles

peab pidama, et erinevaid mudeleid saab võrrelda vaid siis, kui erinevad mudelid on sobitatud sama arv kordi diferentsitud aegridadele [9].

3. Prognoosimine ainult minevikuhindasid kasutades

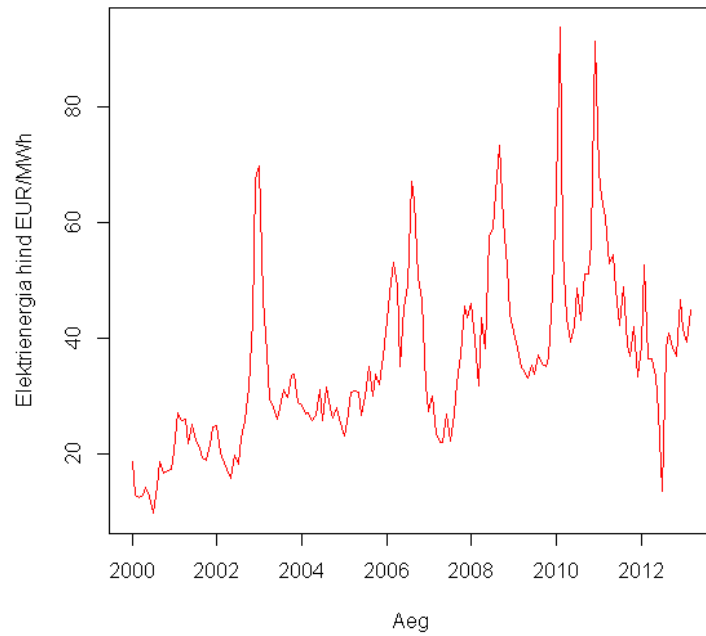
Antud töös on aluseks võetud Soome elektrienergia kuulised hindade andmed, kuna praeguse seisuga on Eesti kohta elektrienergia hindade andmed kätte saadav alates 2010 aasta keskpaigast ning vastavatele andmetele ei ole mõtet leida mudel, kuna see ei pruugi tegelikkust kajastada ning saadud prognoosid võivad olla ebatäpsemad.

Andmestik on pärit *Nord Pool Spot'i* kodulehelt <http://nordpoolspot.com/Market-data1/Elspot/Area-Prices/ALL1/Hourly/>, mis on ühtlasi Euroopa suurim elektriturg. *Nord Pool Spot'i* lehelt on võimalik andmeid vaadelda erinevates valuuta vääringutes – EUR, Norra kroon, SEK. Analüüsis kasutatud andmed on EUR/MWh (LISA 1) ning andmed on ajas kahanevas järjekorras ja kuudele vastavad väärtused on paigutatud tulpadesse. Analüüsimiseks oleks vaja andmed ajas ümber järjestada ajas kasvavaks, st antud juhul esimeseks väärtuseks oleks 2000 aasta jaanuari kuu hinna väärtus ja viimaseks väärtuseks oleks sellisel juhul 2013 märtsikuu hind. Andmetes esineb suur varieeruvus, mis võib kahe järjestikuse kuu vahena küündida 35 euronit MWh kohta. Vaadeldud perioodi maksimaalne väärtus 93,7 oli 2010. veebruaris, mis oli tingitud eriti külmast talvest.

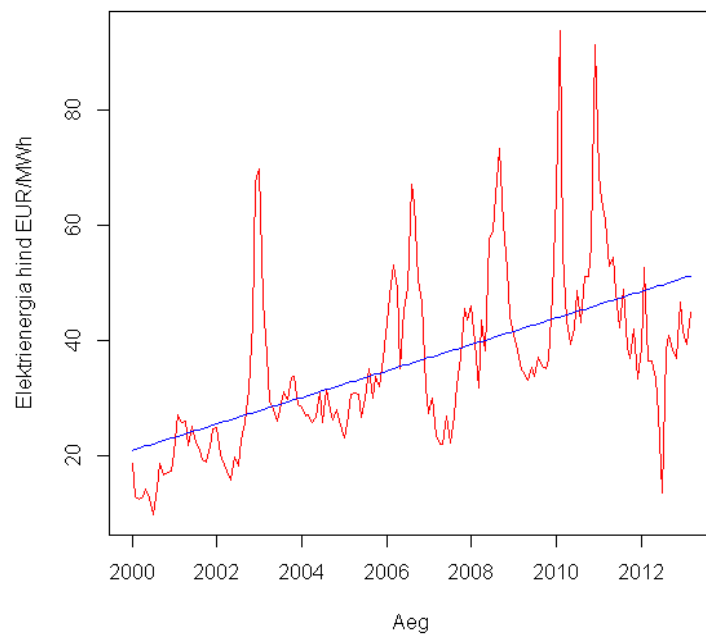
3.1 Soome elektrienergia originaalandmetele sobitatud lihtne mudel

Vaadeldes Soome elektrienergia ajaloolisi hindasid, siis võib täheldada hindade ajas kasvamist. Joonisel 1 paistab, et aegrea väärtuste keskmine on ajas kasvamine, seetõttu ei ole statsionaarne. Kui lineaarse trendi olemasolu on vaadeldava aegrea puhul põhjendatav, tuleb see R-s igal juhul andmetest arima mudeli sobitamisel eemaldada. Edasi tuleb otsustada, kas piisab trendi eemaldamisest või tuleks ka rida diferentsida. Phillips – Perroni (PP) test ütleb, et pigem on tegemist juhuga, kus pärast

lineaarse trendi eemaldamist on meil statsionaarne rida, mistõttu eemaldame andmetest lineaarse trendi.



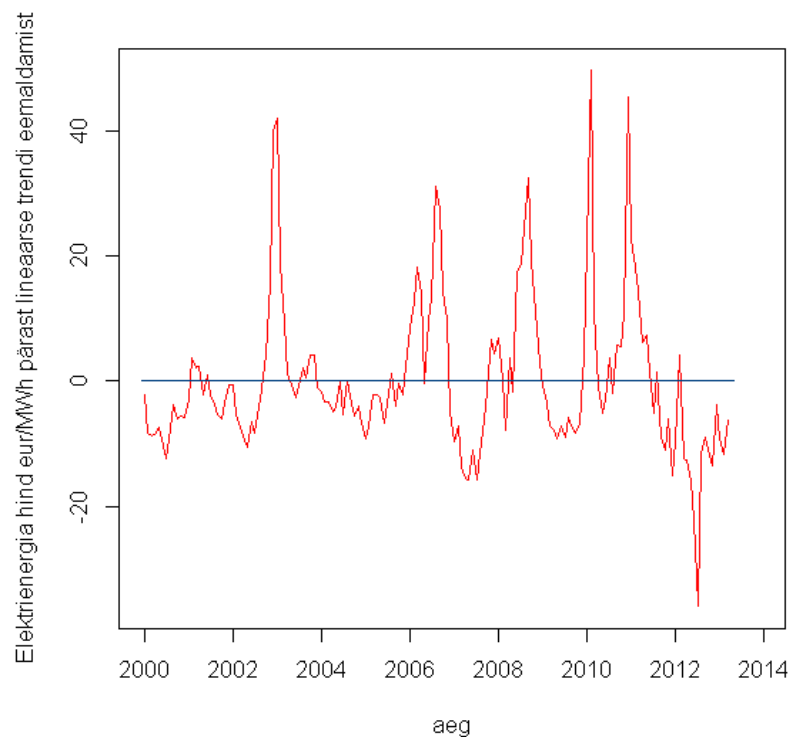
Joonis 1. Soome elektrienergia hinna aegrea graafik



Joonis 2. Soome elektrienergia hinna aegrea graafik koos lineaarse trendiga.

Joonisel 2 on kujutatud Soome elektrienergia hinna aegrea graafik koos lineaarse trendiga, mis on tähistatud joonisel sinise joonega.

Parima mudeli leidmiseks oleks mõistlik esmalt eemalda lineaarne trend ning tulemuseks saaksime uue aegrea graafiku, kus vastava aegrea väärtuste keskvärtuseks on võrdne nulliga. Sellist aegrida võime nimetada ka statsionaarseks aegreaks (vt Joonis 3). Järgneval joonisel on tähistatud saadud aegrida punasega ning sinine joon tähistab keskvärtust.

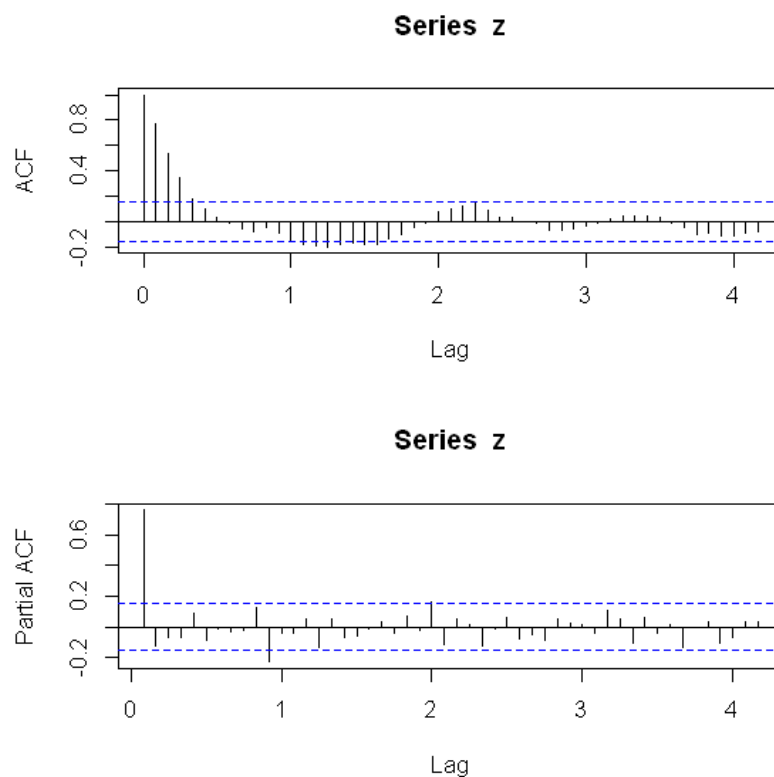


Joonis 3. Elektrienergia hindadele vastav aegrida pärast lineaarse trendi eemaldamist koos keskvärtusega.

Leiame eemaldatud trendiga reale sobiva mudeli. Selleks vaatleme vastava aegrea autokorrelatsioone, lühemalt ACF (*ing Autocorrelation function*), ja osautokorrelatsioone, lühemalt osa ACF (*ing Partial Autocorrelation funkction*), mis on kujutatud Joonisel 4. Siin tuleb mõista, et juhuslikkusest tingituna võime saada andmete pealt hinnates nullist erinevaid autokorrelatsioone ja osautokorrelatsioone

ka juhul, kui rida vastab mudelile, kus tegelikud vastavat järku autokorrelatsioonid või osautokorrelatsioonid on võrdsed nulliga. Et eristada juhuslikkusest tingitud nullist erinevaid väärtuseid tegelikult nullist erinevatest kordajatest, on joonisel kujutatud 95% tõenäosusega kehtivaid veapiire juhuslikele nullist erinevatele kordajatele.

Autokorrelatsioonidest jäävad esimesed 4 tugevalt veapiiridest välja ning alates 13ndast väärtusest on märgata uuesti autokorrelatsioonide kerget kasvu, mis tingib ühtlasi ka väärtuste veapiiridest välja jäämise. ACF graafikult on märgata ka perioodilisust (Joonisel 4 ülemine graafik), kuid see periood ei paista olevat hästi seotud loodusliku, st aastase perioodiga, milleks on meie andmete puhul 12 kuud. Vaadeldes joonise 4 alumist graafikut (*Partial ACF*) paneme tähele, et siin on tugevalt esimene ja natuke ka üheksas väärtus veapiiridest väljas, kuid osa-ACF graafikul puudub korrapärane kõikumine.



Joonis 4. Soome elektrienergia vastava aegrea autokorrelatsioonid ja osautokorrelatsioonid.

Alustame mudeli sobitamist võimalikult väheste parameetrite arvuga mudelitest. Joonise 4 põhjal oleks mõistlik vaadelda algul esimest järku autoregressiivset mudelit ehk AR(1) mudelit. Kui vastav mudel sobib, siis ei ole mõistlik edasi vaadelda rohkemate parameetrite arvuga mudeleid.

Ljung – Box test on statistikas kasutatav test, mis testib kas vastava aegrea autokorrelatsioonide väärtused on nullist erinevad. Testi saab defineerida hüpoteeside kaudu – nullhüpotees H_0 ja alternatiivne hüpotees H_1 .

H_0 : Andmed vastavad sõltumatutele juhuslikele suurustele ning H_1 : Andmed on omavahel sõltuvuses.

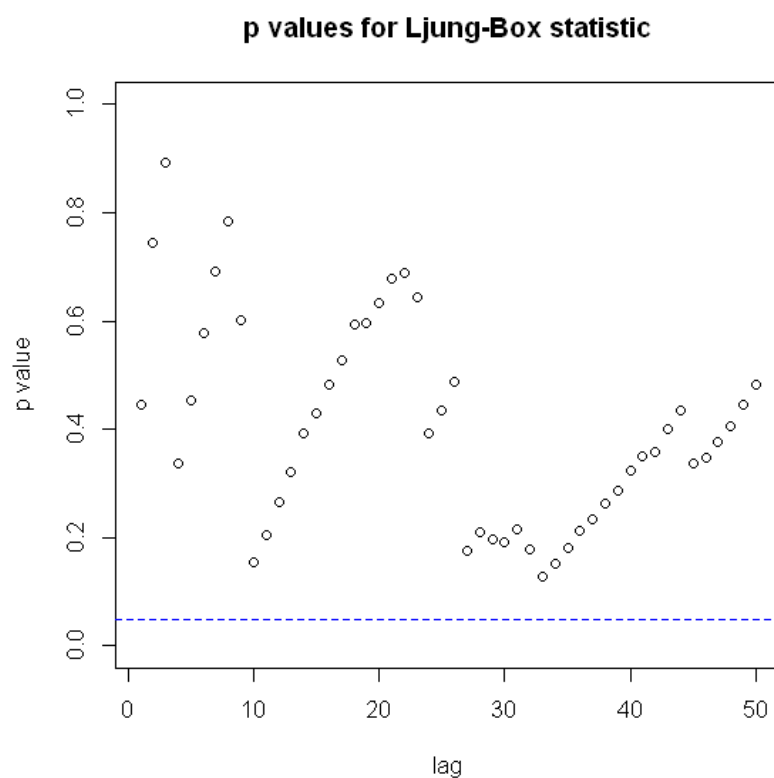
Ljung- Box teststatistiku väärtuse leidmiseks kasutatakse valemit

$$Q = n(n + 2) \sum_{k=1}^n \frac{\hat{\rho}_k^2}{n - k},$$

kus n on valimi suurus, $\hat{\rho}_k^2$ nihkele k vastava autokorrelatsiooni hinnangu ruut ning h tähistab koos vaadeldavate autokorrelatsioonide rühma suurust. Tihti kasutatakse vastavat testi ARIMA mudelite õigsuse kontrollimiseks, mida rakendatakse mudeli jääkidele mitte originaalreale. Sellest tulevalt on nullhüpoteesiks, et mudeli vigade vahel ei ole korrelatsiooni, st lineaarset sõltuvust.

Mudeli sobivust kontrollime Ljung – Box teststatistikuga. Kui kõikidele autokorrelatsioonidel rühmadele suurusega $1, 2, \dots, m$ (kus m vastab paarile-kolmele aastale) vastavad Ljung – Box statistiku p -väärtused on suuremad kui 0.05, siis võib järeldada, et vastav mudel sobib andmetele. Vastasel korral tuleks mudelile lisada juurde mõni lisaparameeter.

Antud juhul sobib meile AR(1) mudel väga hästi – vastavad Ljung – Box väärtused on kindlalt suuremad väärtusest 0.05 (Joonis 5). Sellest tulenevalt ei ole vaja meil vaadelda alternatiivseid mudeleid.



Joonis 5. Mudeli AR(1) Ljung – Box teststatistiku p-väärtuste graafik

Mudeli AR(1) Akaike ja Bayesi Informatsiooni kriteeriumite väärtused on toodud välja Tabelis 1.

Tabel 1

Mudel	Ljung – Box p-väärtused	AIC	BIC
AR(1)	Kõik > 0.05	1117,68	1264.35

3.1.1 Mudeli väljakirjutamine

Soome elektrienergia hindades esines lineaarne trend ning peale selle eemaldamist sobitatud mudelitest sobis andmestikule esimest järku autoregresiivne mudel ehk AR(1) mudel.

Vaadeldes lihtsalt lineaarse trendi mudelit (Joonis 6), siis paneme tähele, et mudelis osutusid vajalikuks kõikide liikmete kordajad, kuna väärtused on suuremad kui vastavad kahekordsed standardhälbed.

Call:

```
arima(x = hind, xreg = aeg)
```

Coefficients:

```
      intercept    aeg  
      20.7958 0.1915  
s.e.    1.9593 0.0212
```

sigma^2 estimated as 151.2: log likelihood=-624.57

AIC=1255.14 AICc=1255.3 BIC=1264.35

Joonis 6. Andmete lineaarse trendi parameetrite määramine

Pärast regressioonijääkidele sobiva mudeli kuju kindlaks tegemist saame R paketi vahendusel sobitada ARIMA mudeli nii, et regressioonisirge kordajad ja mudeli parameetrid leitakse korraga, mistõttu esialgse regressioonisirge kordajad veidi muutuvad. Tulemus on kajastatud Joonisel 7, kus on näha, et parameetrite standardhälbed on kaks korda väiksemad kui vastavate parameetrite väärtused, seega kõik mudeli liikmed on olulised (vt Joonis 7).

Call:

```
arima(x = hind, order = c(1, 0, 0), xreg = aeg)
```

Coefficients:

```
      ar1 intercept    aeg  
      0.7621 20.8535 0.1887  
s.e. 0.0504  5.0395 0.0541
```

sigma^2 estimated as 62.46: log likelihood=-554.74

AIC=1117.47 AICc=1117.73 BIC=1129.75

Joonis 7. Sobiva mudeli parameetrite väärtused

Mudel sai sobitatud reale, millel oli eemaldatud ajaline komponent, st reale

$$Y_t = Z_t - c_1 * t - c_0,$$

kus Z_t on esialgne aegrida, c_1 ja c_0 on koefitsiendid ning t on aeg kuudes.

Kasutades definitsiooni 4 saame mudeli kujuks:

$$(1 - 0.7621B)Y_t = A_t$$

ehk

$$Y_t = 0.7621Y_{t-1} + A_t.$$

Asendame Y_t eeltoodud valemiga saame

$$Z_t - 0.1887t - 20.8535 = 0.7621(Z_{t-1} - 0.1887(t-1) - 20.8535) + A_t.$$

Viime kõik liikmed paremale poole võrdusmärgi, v.a suurus Z_t

$$Z_t = 0.1887 * t + 20.8535 + 0.7621Z_{t-1} - 0.7621 * 0.1887t + 0.7621 * 0.1887 - 0.7621 * 20.8535 + A_t.$$

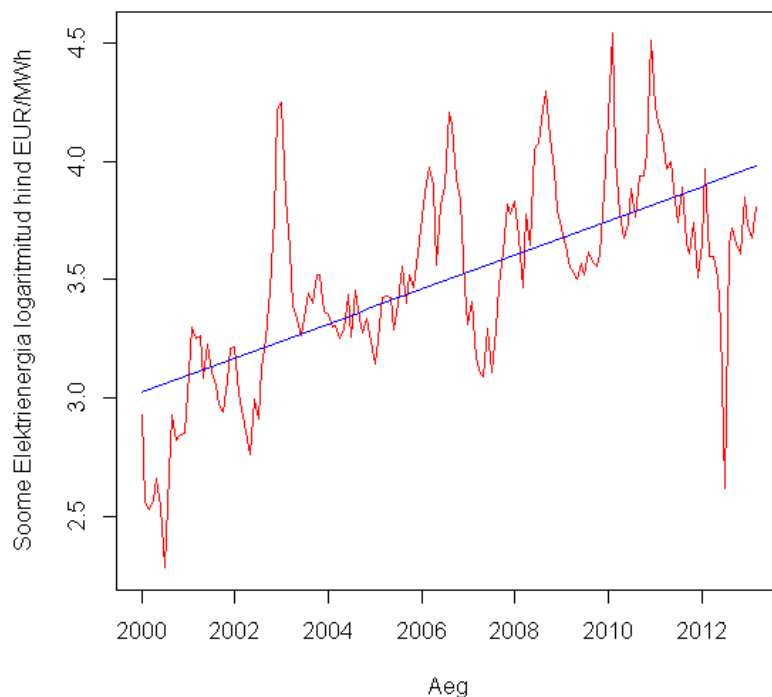
Võttes liikmeti kokku saame tulemuseks mudeli kujul

$$Z_t = 0.7621Z_{t-1} + 0.0526 * t + 5.1049 + A_t.$$

3.2 Elektrienergia hinna logaritmitud andmete lihtne mudel

Joonisel 1 on selgelt näha, et andmete varieeruvus on suur ja esineb ebasümmeetriline kõikumine. Samuti tihtipeale esineb majanduslikes aegridades inflatsioonist tingituna väärtuste ekponentsiaalset kasvumist ajas, sellest lähtuvalt oleks vahel kasulikum vaadelda originaalandmete asemel logaritmitud andmeid (LISA 2). Logaritmimeine vähendab väärtuste vahel suurusi muutusi, st hindade varieeruvus väheneb ning eksponentsiaalse kasvu asemel saadakse lineaarne kasv. Soome elektrienergia logaritmitud hindadest moodustatud aegrida on kujutatud Joonisel 8. Sarnaselt nagu

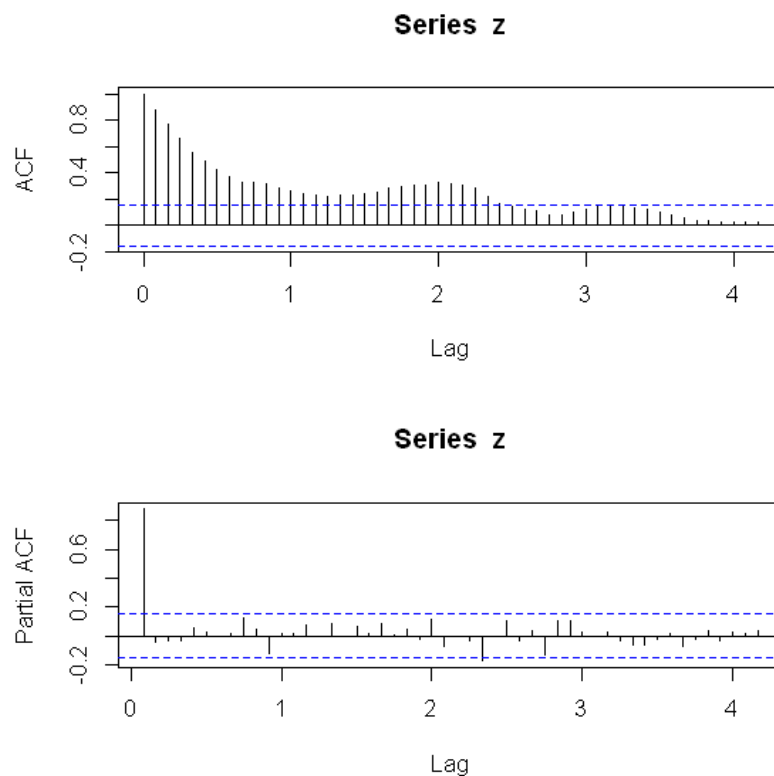
originaalandmete (Joonis 2) puhul on ka siin selgelt eristatav lineaarse trendi olemasolu, mis on kujutatud Joonisel 8 sinise sirgega.



Joonis 8. Soome elektrienergia logaritmitud hindadele vastav aegrida koos lineaarse trendiga.

Sarnaselt punktile 2.2.1 vaatleme logaritmitud andmetega aegrea autokorrelatsioone ja osautokorrelatsioone, mille abil saame kindlaks määrata võimalikud sobivad mudelid, kuid esmalt on mõistlik vajadusel eemaldada lineaarne trend.

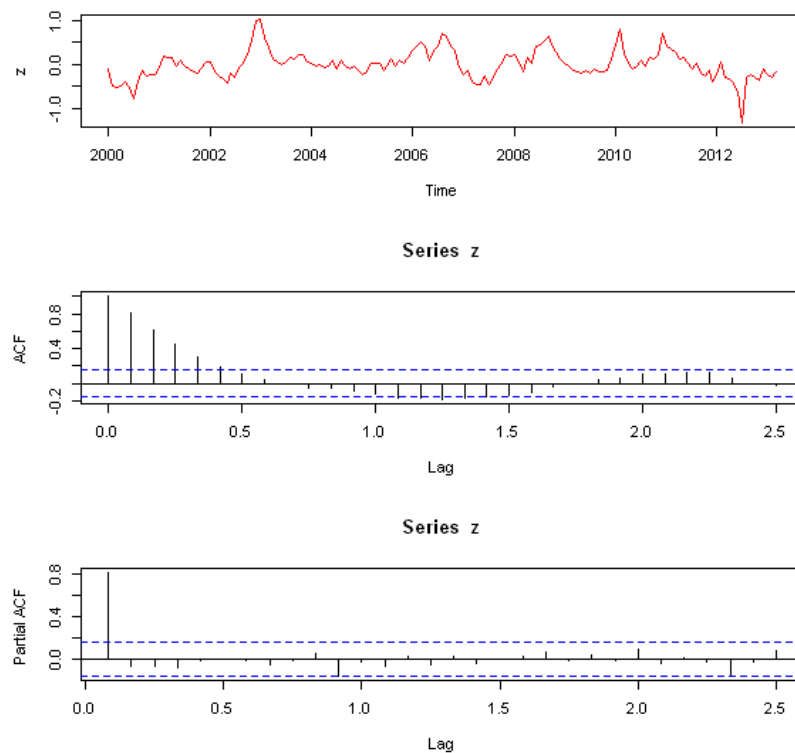
Uurime kõigepealt logaritmitud hinna reale vastavaid autokorrelatsioone ja osautokorrelatsioone. Joonisel 9 on näha, et esimesed kakskümmend üheksa autokorrelatsiooni on selgelt veapiiridest väljas (Joonis 9 *ACF*) ning osautokorrelatsioonidest (Joonis 9 *Partial ACF*) on madalamas järgus ainult esimene veapiirist väljas. Philips – Perroni testi tulemuse alusel ei ole antud aegrida mõistlik diferentsida.



Joonis 9. Logaritmitud andmete autokorrelatsioonid ja osautokorrelatsioonid.

Eemaldame logaritmitud andmetest lineaarse trendi ning vaatleme saadud aegrea autokorrelatsioone ja osautokorrelatsioone. Tulemus on Joonisel 10, kus on näha, et saadud rida on statsionaarne, kuna vastava rea autokorrelatsioonid kahanevad kiirelt nulli. Joonise 10 ACF graafikult on näha, et esimesed viis autokorrelastiooni jäävad veapiiridest välja. Samas Joonise 10 *Partial ACF* graafikul on ainult esimene osautokorrelatsioon, mis jääb veapiirist välja. Perioodile vastavas osas jäävad autokorrelatsioonid natuke veapiiridest välja.

Kuna eesmärk on leida võimalikult väheste kordajatega mudelit, siis esimeseks vaatlusaluseks mudeliks on mõstlik võtta jällegi esimest järku autoregresiivne mudel, st AR(1) mudel.



Joonis 10. Logaritmitud andmete aegrea, ACF ja osa-ACF graafikud peale lineaarse trendi eemaldamist

Saadud tulemus on välja toodud Tabelis 2, kus esimeses tulbas on mudel, teises Ljung – Box statistiku p-väärtuste võrdlus veapiiriga ning viimastes tulpades AIC ja BIC väärtused.

Tabel 2

Mudel	Ljung – Box väärtused	AIC väärtus	BIC väärtus
AR(1)	Kõik > 0.05	-67.87	-55.6

Tulemuseks saime, et vaadeldud mudel sobib meie andmestikule hästi. Alternatiivseid mudelid ei ole edasi mõtet vaadelda, kuna eesmärk leida vähimate kordajatega sobiv mudel on täidetud. Etteruttavalt võib öelda, et saadud AIC väärtust ei saa varasemas punktis leitud parima mudeli AIC väärtusega üks-ühele võrrelda, kuna mudelid on

sobitatud erinevatele andmetele (esimesel juhul originaalandmed, teisel juhul logaritmitud andmed).

3.2.1 Logaritmitud andmete mudeli väljakirjutamine

Eelmises punktis osutus parimaks mudeliks esimest järku autoregressiivne mudel ehk AR(1) mudel, millelt oli eemaldatud lineaarne trend. Enne mudeli väljakirjutamist vaatleme lineaarse trendi parameetrite väärtusi, mis on välja toodud Joonisel 11, kus esimene rida tähistab kordajate väärtusi ja teine rida vastavate väärtuste standardhälbeid.

Call:

```
arima(x = loghind, xreg = aeg)
```

Coefficients:

```
      intercept  aeg  
      3.0164 6e-03  
s.e.  0.0508 6e-04
```

```
sigma^2 estimated as 0.1017: log likelihood=-43.93  
AIC=93.86 AICc=94.02 BIC=103.07
```

Joonis 11: Logaritmitud andmete lineaarse trendi parameetrite väärtused

Mudeli väljakirjutamisel tuleb lugeda olulisteks kõik mudeli parameetrite väärtused, kuna kõik on suuremad oma kahekordsest standardhälbest.

Call:

```
arima(x = loghind, order = c(1, 0, 0), xreg = aeg)
```

Coefficients:

```
      ar1 intercept  aeg  
      0.7993  3.0153 0.0060  
s.e.  0.0466  0.1420 0.0015
```

```
sigma^2 estimated as 0.0361: log likelihood=37.94  
AIC=-67.87 AICc=-67.61 BIC=-55.6
```

Joonis 12: Logaritmitud andmete AR(1) mudeli parameetrite väärtused
 Kasutades programmi R saame logaritmitud andmetele sobitada ARIMA tüüpi mudelit nii, et ühtlasi leitakse ka sobiv lineaarne trend, mis lähteandmetest eemaldatakse. Tulemus on Joonisel 12, kus on selgelt näha, et antud juhul on meil kõik liikmed olulised.

Kirjutame logaritmitud aegreale vastava mudeli välja, saame

$$(1 - 0.7993B)Y_t = A_t,$$

kus B tähistab endiselt tagasinihke operaatorit ning viies mudeli meile sobivamale kujule, saame tulemuseks mudeli kujul

$$Y_t = 0.7993Y_{t-1} + A_t.$$

Asendades $Y_t = \ln(Z_t) - c_1 * t - c_0$ saame

$$\ln(Z_t) - 0.006 * t - 3.0153 = 0.7993(\ln(Z_{t-1}) - 0.006 * (t - 1) - 3.0153) + A_t$$

ehk

$$\ln(Z_t) - 0.006 * t - 3.0153 = 0.7993 \ln(Z_{t-1}) - 0.0048 * t - 2.4053 + A_t.$$

Peale lihtsustamist saame

$$\ln(Z_t) = 0.7993 \ln(Z_{t-1}) + 0.0012 * t + 0.61 + A_t.$$

Võtame mõlemad pooled e astmesse ning saame tulemuseks

$$e^{\ln(Z_t)} = Z_t = e^{0.7993 \ln(Z_{t-1}) + 0.0012 * t + 0.61 + A_t}$$

ehk

$$Z_t = e^{0.7993 \ln(Z_{t-1}) + 0.0012 * t + 0.61} e^{A_t}. \quad (1)$$

Selle mudeli põhjal prognooside arvutamiseks vaatleme üldisemalt prognoosimist juhul, kui mudel on sobitatud logaritmitud reale.

Vaatleme aegrida $Y_t = \ln(Z_t)$. Oletame, et Y_t vastab mingile ARIMA tüüpi mudelile ning olgu \hat{Y}_t mudeli põhjal prognoositud Y_t väärtus, siis

$$Y_t = \hat{Y}_t + A_t,$$

kus A_t on sõltumatud ajal $t - 1$ olemasolevast infost.

Võrdusest

$$Z_t = e^{Y_t}$$

saame nüüd Z_t prognoosiks

$$E(Z_t | Z_{t-i}, i = 1, 2, \dots) = e^{\hat{Y}_t} * E(e^{A_t}).$$

Eeldame, et A_t on normaaljaotusega $N(0, \sigma)$, siis e^{A_t} on lognormaalse jaotusega.

Kasutades teadmist, et jaotusega $LN(\mu, \sigma)$ juhusliku suuruse keskväärtus on $e^{\mu + \frac{\sigma^2}{2}}$, siis

$$E(A_t) = E(\ln e^{A_t}) = e^{\frac{\sigma^2}{2}}.$$

Seega saame

$$E(Z_t | Z_{t-i}, i = 1, 2, \dots) = e^{\hat{Y}_t} * E(e^{A_t}) = e^{\hat{Y}_t + \frac{\sigma^2}{2}}.$$

Asendades mudelis (1) väärtus $A_t = 0.0361$ (Joonisel 12 σ^2 estimated as), saame prognooside mudeli kujuks

$$Z_t = e^{0.7993 \ln(Z_{t-1}) + 0.0012 * t + 0.61} e^{0.0361}.$$

3.3 Lihtsate mudelite võrdlus ja lühiajaline ennustamine

Punktides 2.2.1.1 ja 2.2.2.1 saime kaks erinevat mudeli kuju, mis mõlemad sobivad Soome elektrienergia andmetele. Erinevus tuleneb sellest, et 2.2.1.1 mudel on saadud algandmete põhjal ning punktis 2.2.2.1 logaritmitud andmete põhjal, mistõttu saadud prognoose tuleb teisel juhul teisendada prognoosideks esialgse aegrea jaoks.

Erinevate mudelite võrdlemiseks on mitu erinevat viisi, millest tuntumad ja kasutatavamad on keskmine absoluutne viga (MAD), keskmine ruutviga (MSE), ruutkeskmine viga (RSME) ja keskmine suhteline viga (MAPE). Samuti kasutatakse

Akaike informatsiooni kriteeriumit ja Bayesi Informatsiooni kriteeriumit (BIC), kuid viimast on kasutatav juhul, kui mudelid on sobitatud samadele andmetele.

Leiame prognoosivigade mõõdikud nii originaalandmete jaoks leitud mudeli (tähistus $AR(1)$) kui ka logaritmitud andmete jaoks leitud mudeli (tähistatud $AR(1)_{Log}$). Saadud väärtused on toodud Tabelis 3.

Tabel 3

Mõõdikud	Mudel	$AR(1)$	$AR(1)_{Log}$
MAD		5.31679	5.790295
MSE		62.45554	62.62068
RMSE		7.902882	7.913323
MAPE		0.1471742	0.1475768
AIC		1117.73	
BIC		1129.75	

Tabelis 3 saadud väärtuste põhjal saab öelda, et kahest leitud mudelist osutus parimaks originaalandmetele sobitatud mudel, kuna Tabelis 3 toodud vastava mudeli mõõdikute väärtused on väiksemad võrreldes logaritmitud andmetele sobitatud $AR(1)$ mudeliga.

Elektrienergia hinna ennustamisel loetakse lühiajaliseks perioodiks, mis jääb vahemikku üks kuu kuni kaksteist kuud. Antud juhul prognoosime 9 kuu väärtusi, st käesoleva aasta lõpuni. Prognoosimisel võtame aluseks parimaks mudeliks osutunud mudel $AR(1)$ (mudel kujul (1)) ning vastav tulemus on välja toodud Tabelis 4. Soome elektrienergia prognoositud hinnad on ajas kasvavad ning järjestikuste väärtuste vaheline samm ei ole ühtlane.

Tabel 4

Aeg	Elektrienergia hinna prognoositud väärtus
Aprill 2013	46.59107
Mai 2013	47.84087
Juuni 2013	48.83822
Juuli 2013	49.64317
August 2013	50.30151
September 2013	50.84812
Oktoober 2013	51.30957
November 2013	51.70614
Detsember 2013	52.05326

Leiame Soome elektrienergia prognoositud hindade veapiirid 95% ja 80% usaldusnivool, mis on väljatoodud Tabelis 5.

Tabel 5

Aeg	Alumine veapiir		Ülemine veapiir	
	80%	95%	95%	80%
Aprill	36.47539	31.10143	62.08072	56.70676
Mai	35.12261	28.36603	67.31572	60.55914
Juuni	34.82869	27.38612	70.29031	62.84775
Juuli	34.93568	27.12232	72.16402	64.35066
August	35.20347	27.18264	73.42038	65.39955
September	35.52784	27.38894	74.30729	66.16839
Oktoober	35.86170	27.65501	74.96414	66.75745
November	36.18464	27.93884	75.47344	67.22765
Detsember	36.48917	28.22074	75.88579	67.61736

4. Prognoosimine täiendavaid regressoreid kasutades

Eelmises osas vaatlesime Soome elektrienergia hinna kõige lihtsamat mudelit, milleks osutus AR tüüp mudel (1). Selles osas vaatleme, kas keskkonna mõjutegurid nagu temperatuur ja sademete hulk mõjutavad elektrienergia hinda või mitte. Selleks kasutame ARIMAX mudelit (definiitsioon 6). Selleks tuleb meil leida ülekandefunktsiooni $\beta(x)$ kordajate hinnangud, kuna regressori X omandatud väärtused kanduvad üle Z väärtustele. Mudeli leidmine on mitme etapiline:

1. Suurustele β_i hinnangute leidmine
2. Nihke b leidmine ja hinnangute leidmine parameetritele r ja s
3. Vigadele ε_t sobiva mudeli leidmine
4. Mudeli parameetrite hindamine suurima tõepära meetodil
5. Jääkvigade sõltumatuse kontrollimine

Suuruste β_i hindamisel kasutatakse nn eelvalgendamise tehnikat (Ing *prewhitening*), mis on kasutatav juhul, kui protsess X vastab mingile pööratavale ARMA tüüpi mudelile. Eelvalgendamine on mitmesammuline:

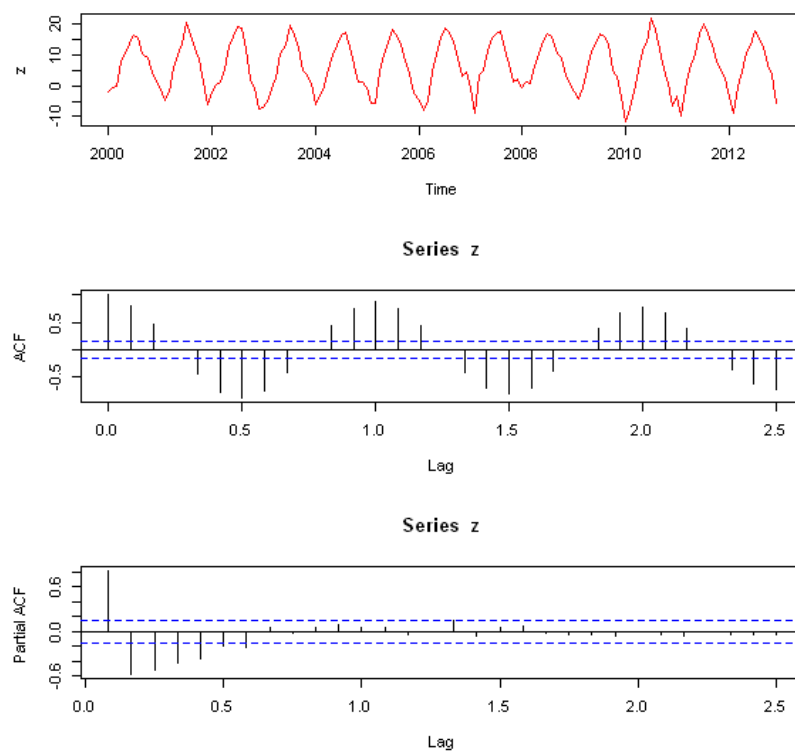
1. Määrame regressor rea mudeli ja leiame vastava mudeli jäägid. Siin ei ole eemärk leida parim mudel, vaid mudel, mis sobib ning liikmete arv oleks mõistlik. Regressor rida X võib tõlgendada kui „valget müra“.
2. Sobitame eelmises punktis leitud mudeli kuju Y protsessile ja leiame mudeli jäägid.
3. Leiame ristkorrelatsioonid kahe eelneva punktis leitud mudeli jääkide vahel.

Tihti kasutatakse esimeses punktis sobitamisel $AR(p)$ tüüpi mudeleid, kuna nende hilisem kasutamine on lihtsam [17].

4.1 Temperatuuri mõju Soome elektrienergia hindadele

Esmalt vaatlen temperatuuri mõju Soome elektrienergia hindadele. Andmed on saadud EMHI (Eesti Metroloogia ja Hüdroloogia Instituut) poolt. Algsed andmed on ühepäevase intervalliga ja vastavad ajaperioodile 2000 aastast kuni 2012 aasta lõpuni. Esmane andmete töötlus on tehtud programmiga *Microsoft Excel*, kus on leitud kuude keskmised temperatuurid kasutades käsku *average* (LISA 3).

Definitsioonis 7 on öeldud, et vaadeldavad aegread peavad olema statsionaarsed, siis sellest tulenevalt peame vaatleme, kas kuude keskmistest temperatuuridest moodustatud aegrida on statsionaarne või esineb perioodilist kõikumist. Vaadeldes aegrea graafikut (Joonis 13 z graafik), siis seal on selgelt eristatav perioodiline kõikumine, milleks on 12 kuud. See nähtus on loomulik, sest see on tugevalt seotud nelja erineva aastaajaga. Talvel temperatuurid üldjuhul negatiivsed, kevadel ja sügisel nulli ja ± 5 kraadi vahel ning suvel rohkem kui viis kraadi Celsiuse järgi.



Joonis 13. Temperatuuri väärtustest moodustatud aegrea, autokorrelatsioonide ja osaaautokorrelatsioonide graafik

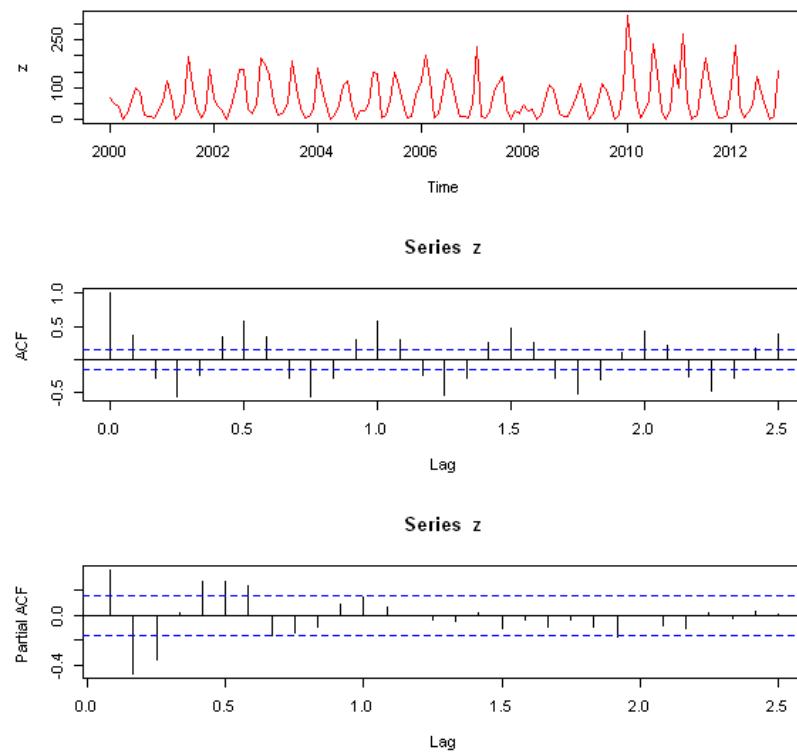
Majanduslikult ei ole mõistlik antud aegrida elektrihinna regressorina kasutada, kuna nii väga madalad kui ka väga kõrged temperatuurid üldiselt tõstavad elektrihinna, kuid temperatuurile vastav liige mõjutaks väga soojade ilmade korral elektrihindu vastupidiselt sellele, milline oleks mõju väga külmade ilmade korral. Teisisõnu, kui näiteks kuumad ilmad leitud mudeli põhjal suurendaks elektrihindu, siis külmad ilmad kahandaks hindu ja vastupidi. Selline käitumine ei ole reaalne ega otstarbekas. Kuna nii väga soojad ja väga külmad ilmad suurendavad elektritarbimist, siis on otstarbekas kasutada regressoritena vastava omadustega funktsioone temperatuurist. Käesoleva töös vaatleme kolme võimalikku sobivalt käituvat funktsiooni:

1. $(T_t - \mu)^2$, kus T_t tähistab temperatuuri ajahetkel t ja μ aastast keskmist temperatuuri
2. $|T_t - \mu|$, kus T_t tähistab temperatuuri ajahetkel t ja μ aastast keskmist temperatuuri
3.
$$\begin{cases} T_t - 12, & \text{kui } T_t > 12 \\ 0, & \text{kui } 5 < T_t < 12 \\ 5 - T_t, & \text{kui } T_t < 5 \end{cases}$$

4.1.1 Temperatuuri funktsiooni $(T_t - \mu)^2$ kasutamine regressor reana

Vaatleme juhtu, kus eemaldame temperatuuri väärtustest vaadeldud perioodi keskmise temperatuuri 6.394348 ning võtame saadud vahe ruutu. Tulemuseks saame aegrea, mille kõik väärtused on positiivsed (Joonis 14 z graafik). Vaadeldes antud aegrea autokorrelatsioone (Joonis 14 ACF graafik) ja osaaautokorrelatsioone (Joonis 14 osa-ACF) näeme, et perioodilisus on mingil määral olemas.

Aegrida ei ole mõistlik ennem regressorina kasutamist sesoonselt diferentsida, kuna perioodile vastavad autokorrelatsioonid kahanevad ning elektrihinna jaoks ei ole oluline kahe järjestikuse aasta temperatuuride vaheline erinevus (st kui palju oli sel aastal külmem kui eelmisel aastal), vaid pigem ikkagi konkreetse kuu keskmine temperatuur (kui sellel aastal oli külm veebruar, siis see suurendab vastava kuu elektrihinna sõltumata sellest, milline oli eelmise aasta veebruar).



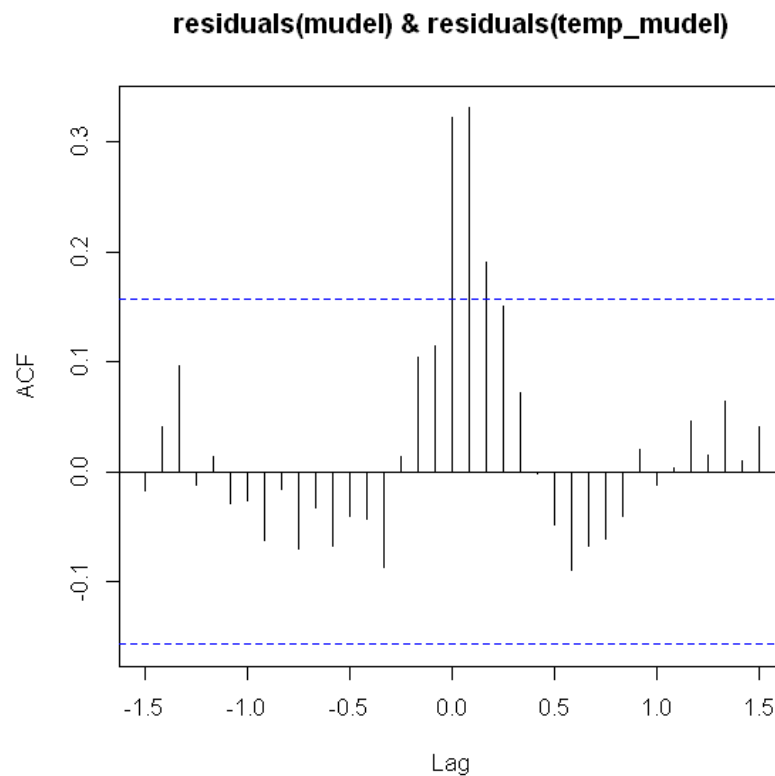
Joonis 14. Keskmise temperatuuri väärtuste aegrida, kus on eemaldatud kogu perioodi keskmine ning vastava aegrea autokorrelatsioonid ja osautokorrelatsioonide graafikud

Esmalt sobitame AR(7) mudelit, kuna osautokorrelatsiooni graafikut vaadates on seitse esimest veapiiridest väljas.

Tabel 6

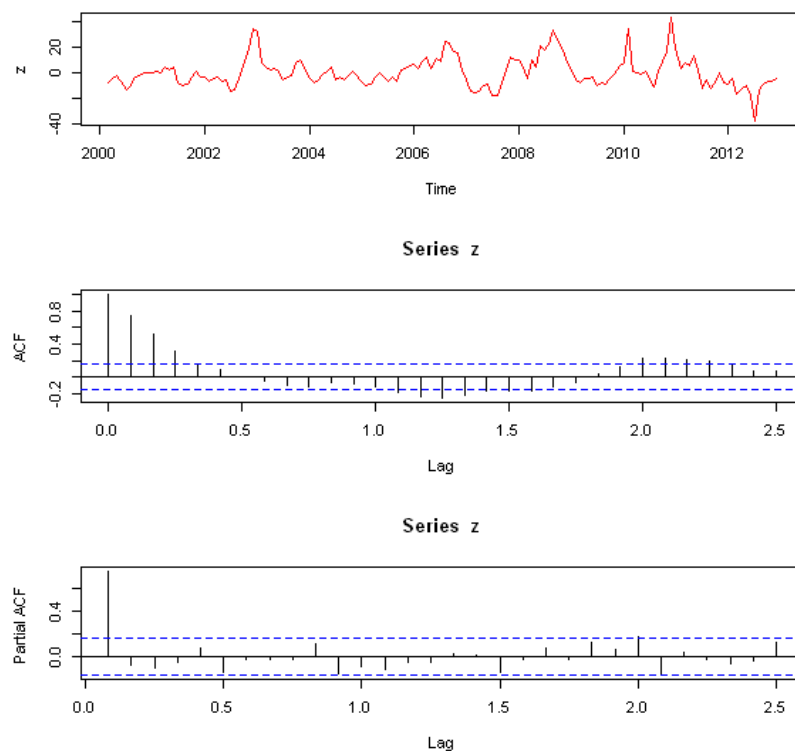
Mudel	Ljung – Box väärtused	AIC väärtus	BIC väärtus
AR(7)	Kõik > 0.05	1575.76	1603.21

Tabelis 6 näeme, et vastav mudel sobib temperatuuri aegreale. Seetõttu ei vaatle edasi alternatiivseid mudeleid. Sobitame AR(7) mudelit nüüd elektrihindade aegreale, millelt on eemaldatud eelnevalt trend ja vaatleme vastavate mudeli jääkide ristkorrelatsioone.



Joonis 15. Mudeli $(T_t - \mu)^2$ ja elektrienergia hinna mudeli jääkide vaheline ristkorrelatsioon

Paneme jooniselt 15 tähele, et nullis, esimene ja teine ristkorrelatsioon on veapiiridest väljas ja seega on meil võimalik ajalooliste temperatuuri väärtustega elektrienergia hindu prognoosida. Rakendades algele aegreale neli regressorit, milleks on vaadeldava temperatuuri funktsiooni aegrida nihkega 0, 1, 2 ning lisaks lineaarsele trendile vastav regressor. Regressorite sobitamise tulemusena tekkinud elektrienergia hinna ennustamise jääkide autokorrelatsioonid ja osautokorrelatsioonid on toodud Joonise 16.



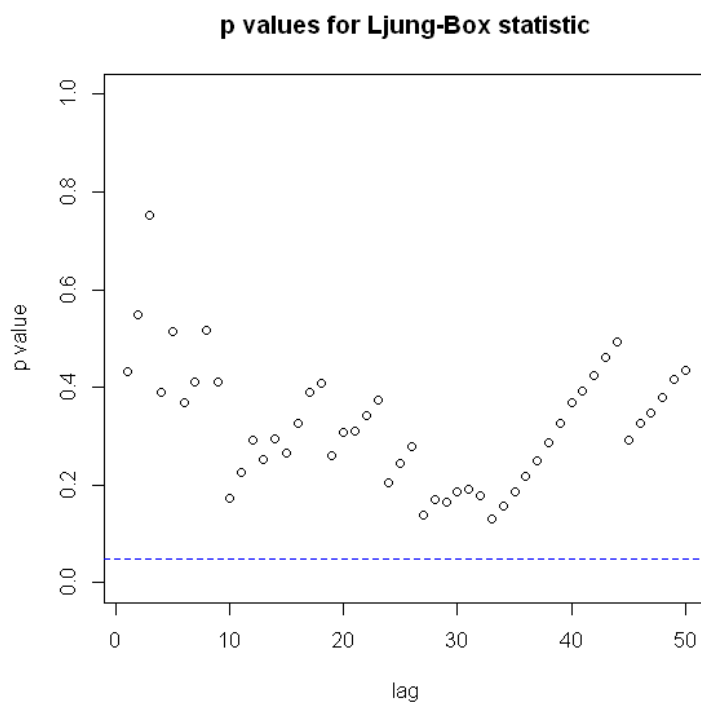
Joonis 16. Elektrienergia hinna aegrea, ACF ja osa- ACF graafikud kasutades temperatuuri kui regressorrida

Sellise joonise põhjal on esimeseks mudeliks $AR(1)$, kuna ainult esimene osautokorrelatsioon on veapiirist väljas. Kui esmane mudel ei sobi, siis alternatiivseks mudeliteks võib vaadelda sega mudelit $ARMA(1,0,1)$.

Tabel 7

Mudel	Ljung – Box väärtused	AIC väärtus	BIC väärtus
ARX(1)	Kõik > 0.05	1067.27	1088.53

Tulemuseks saime, et mudel $AR(1)$ sobib, kuna Ljung – Box teststatistiku p-väärtused on suuremad kui 0.05 (Joonis 17) ja seetõttu ei vaatle ka alternatiivset mudelit.



Joonis 17. ARX(3) Ljung – Box p-väärtused

Uurime, kas kõik regressorid on vajalikud elektrienergia hindade prognoosimiseks. Võrdleme Joonise 18 parameetrite väärtusi vastavate kahekordse standardhälvetega. Selgub, et nihkele kaks vastav regressor ei ole oluline ja seega võime välja jätta.

Series: z
 ARIMA(1,0,0) with non-zero mean

Coefficients:

	ar1	intercept	xreg1	xreg2	xreg3	xreg4
	0.7665	16.5852	-0.0019	0.0315	0.0337	0.1916
s.e.	0.0511	5.0503	0.0096	0.0091	0.0094	0.0538

sigma^2 estimated as 54.37: log likelihood=-526.63
 AIC=1067.27 AICc=1068.03 BIC=1088.53

Joonis 18. ARIMAX mudeli parameetrite kordajate väärtused

Kokkuvõttes saame mudeli ARX(1), mille regressoriteks on temperatuur nihketega 1, 2 ja regressioonisirge (Joonis 19).

Series: z

ARIMA(1,0,0) with non-zero mean

Coefficients:

	ar1	intercept	xreg2	xreg3	xreg4
	0.7657	16.4402	0.0314	0.0343	0.1914
s.e.	0.0511	4.9853	0.0091	0.0090	0.0536

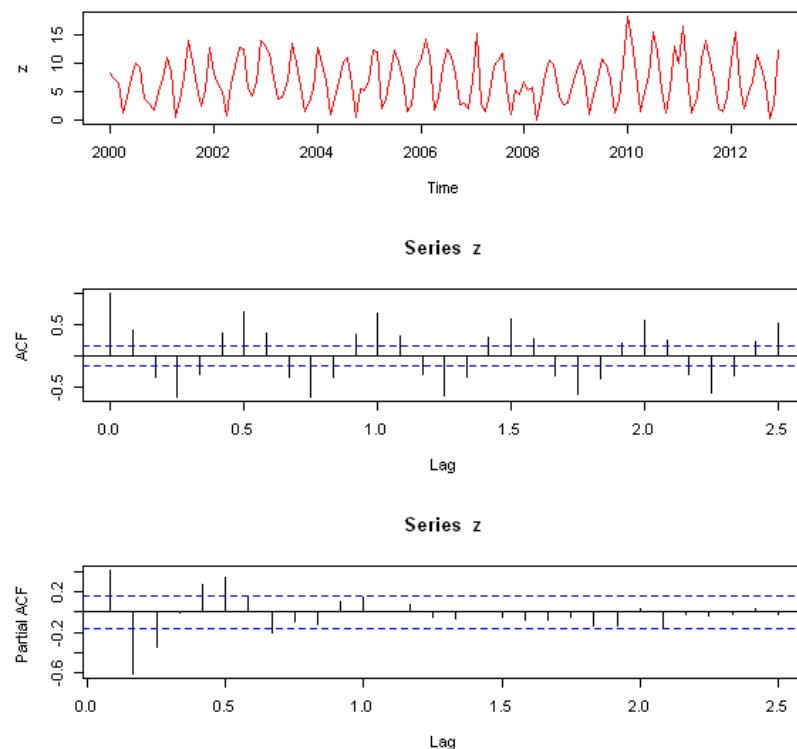
sigma^2 estimated as 54.38: log likelihood=-526.65

AIC=1065.31 AICc=1065.88 BIC=1083.53

Joonis 19. ARIMAX parameetrite kordajate väärtused

4.1.2 Temperatuuri funktsiooni $|T_t - \mu|$ kasutamine regressor reana

Teisena kirjeldatud funktsioon, kus igal ajahetke on temperatuurist maha lahutatud temperatuuri keskmine väärtus μ ja seejärel vahest võetud absoluutväärtus. Vaadeldes Joonist 20, siis paneme tähele, et nii autokorrelatsioonid kui ka osautokorrelatsioonid käituvad eelmise funktsiooni aegreaga (Joonis 14) suhteliselt sarnaselt.



Joonis 20. Teise funktsiooni aegrida, ACF ja osa-ACF graafikud

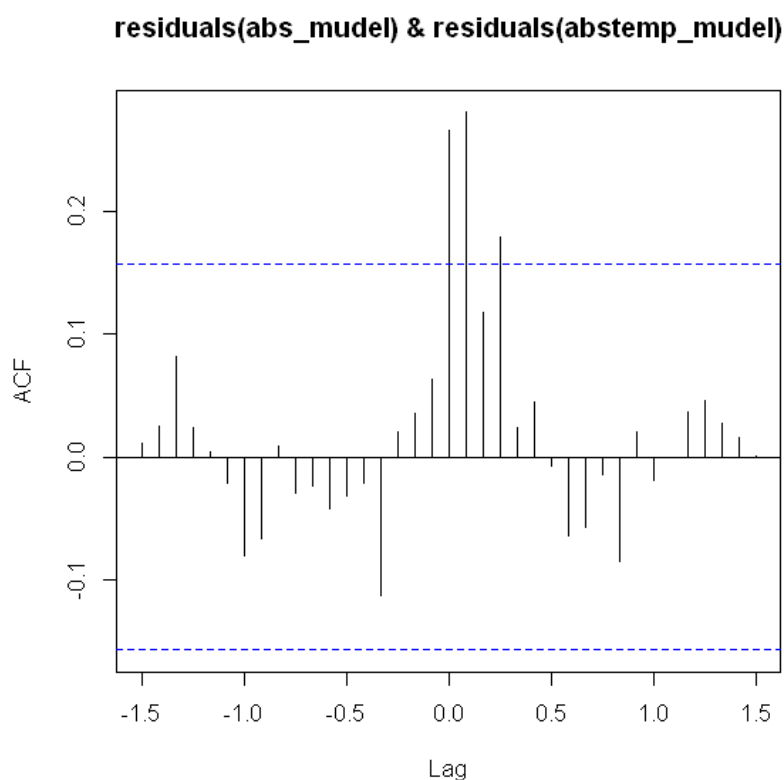
Ka siin ei ole mõistlik ennem aegrida sesoonselt diferentsida, kuna perioodile vastavad autokorrelatsioonid kahanevad. Proovime sobitada vaadeldavale aegreale ARMA tüüpi mudelit võttes aluseks osautokorrelatsioonide graafiku (Joonis 20 *Partial ACF*). Esmalt sobitame AR(8), kuna osautokorrelatsioonide graafikut vaadates on esimesed kaheksa väärtust veapiiridest väljas.

Tabel 8

Mudel	Ljung – Box väärtused	AIC väärtus	BIC väärtus
AR(8)	Kõik > 0.05	1239.25	1269.75

Tabelis 8 saime, et vastav mudel sobib andmetele ning sellest lähtuvalt ei ole meil mõtet vaadelda teisi mudeleid. Sobitame samasugust, st AR(8) mudelit elektrihinna aegreale, millelt on eelnevalt lineaarne trend eemaldatud.

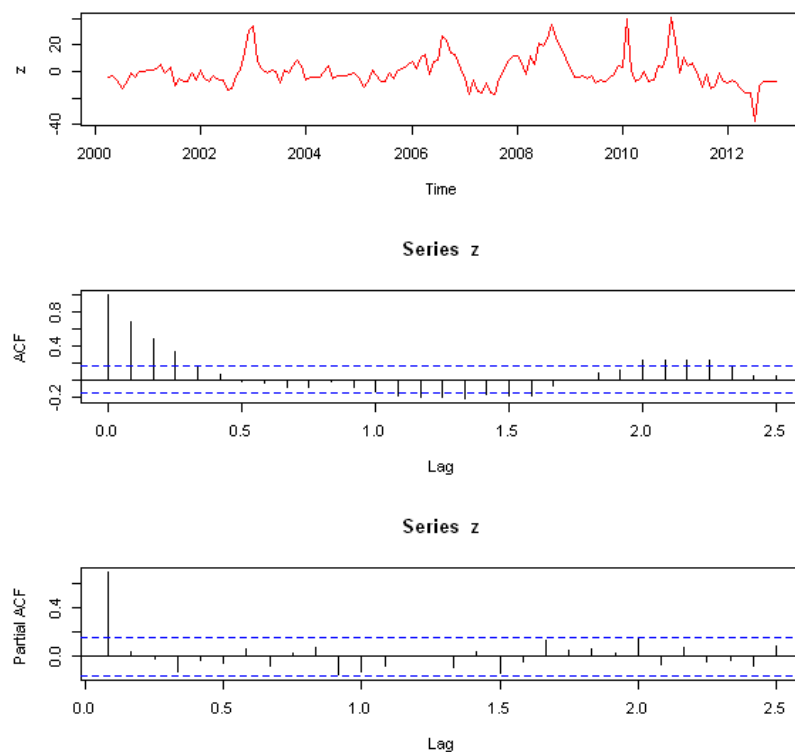
Vaatleme vastavate mudelite jääkide ristkorrelatsioone (Joonis 21), kus on näha, et nullis, esimene ja kolmas ristkorrelatsioon on veapiiridest väljas.



Joonis 21. $|T_t - \mu|$ funktsiooni aegrea ja lineaarse trendita hinna aegrea mudeli jääkide vaheline ristkorrelatsioon

Joonise 21 põhjal saame järeldada, et ajalooliste temperatuuride väärtustest moodustatud funktsiooni väärtustega on võimalik Soome elektrienergia hindade prognoosimisel kasutada. Seega on meil võimalik sobitada ARIMAX tüüpi mudelit.

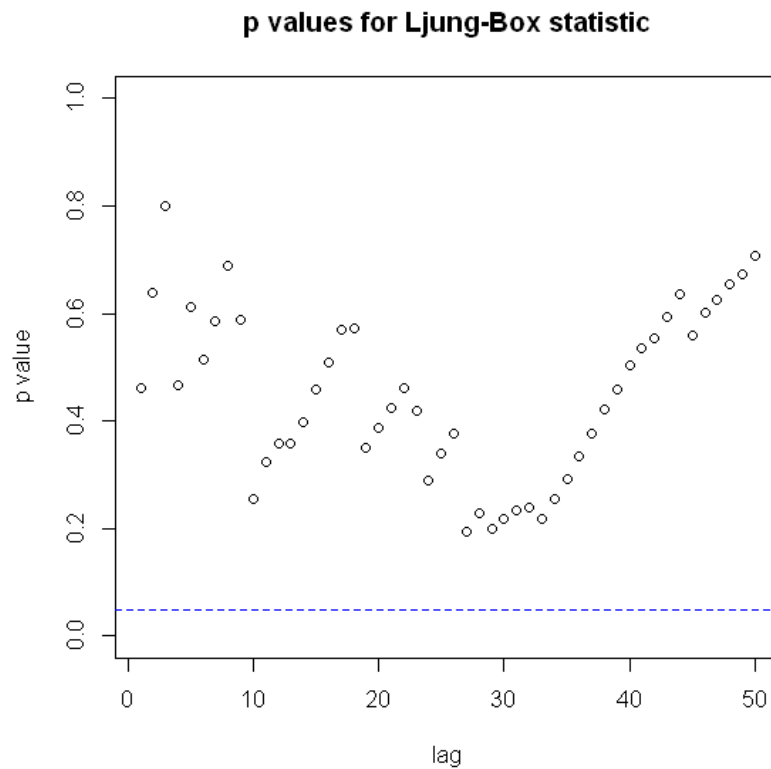
Antud juhul vaatleme temperatuuri aegrida nihetega 0,1 ja 3 ning lisaks veel lineaarsele trendile vastavat regressorit. Sellest tulevalt jätame elektrihindade aegreast välja esimesed kolm väärtust, kuna vastavate väärtuste jaoks pole regressoreid. Joonisel 22 on saadud ilma lineaarse trendita hinna aegrea graafik, millele on võetud arvesse temperatuuri mõju.



Joonis 22. Hinna aegrida koos temperatuuri ja lineaarse trendi mõjuga

Vaadeldes Joonisel 22 autokorrelatsioonide ja osautokorrelatsioonide graafikut, siis esimeseks ARIMAX tüüpi mudeliks võiks sobida AR(1) mudel, kuna ainult esimene

osautokorrelatsioon on väljas. Alternatiivseks mudeliks võib vaadelda segamudelit, kuna paistab, et autokorrelatsioonid kahanevad nulli mingi arvu kordselt. Ljung – Box testi põhjal (Joonis 23) võime lugeda mudeli sobivaks.



Joonis 23. Ljung- Box teststatistiku p-väärtuste graafik

Kuna Joonisel 21 oli kolmas ristkorrelatsioon võrreldes teistega suhteliselt vähe üle veapiiri, siis võrdleme vastavate liikmete kordajate väärtusi vastava väärtuse kahekordse standardhälbega.

Series: z
ARIMA(1,0,0) with non-zero mean

Coefficients:

	ar1	intercept	xreg1	xreg2	xreg3	xreg4
	0.7627	18.8646	-0.0022	-0.0007	0.0340	0.1902
s.e.	0.0523	5.3080	0.0110	0.0101	0.0114	0.0557

sigma^2 estimated as 58.86: log likelihood=-529.28
AIC=1072.57 AICc=1073.34 BIC=1093.78

Joonis 24. ARIMAX parameetrite väärtused

Joonisel 24 on näha, et ebaolulisteks liikmeteks osutus xreg1 ja xreg2 ehk temperatuuri aegrida nihetega 1 ja 3. Vastavad liikmed välja jättes saame veidi teiste väärtustega kordajad (Joonis 25).

```
Series: z
ARIMA(1,0,0) with non-zero mean

Coefficients:
    ar1 intercept xreg2 xreg3 xreg4
    0.7615  18.6595 -4e-04 0.0352 0.1897
s.e. 0.0521   5.1861  1e-02 0.0099 0.0554

sigma^2 estimated as 58.88: log likelihood=-529.31
AIC=1070.61 AICc=1071.19 BIC=1088.79
```

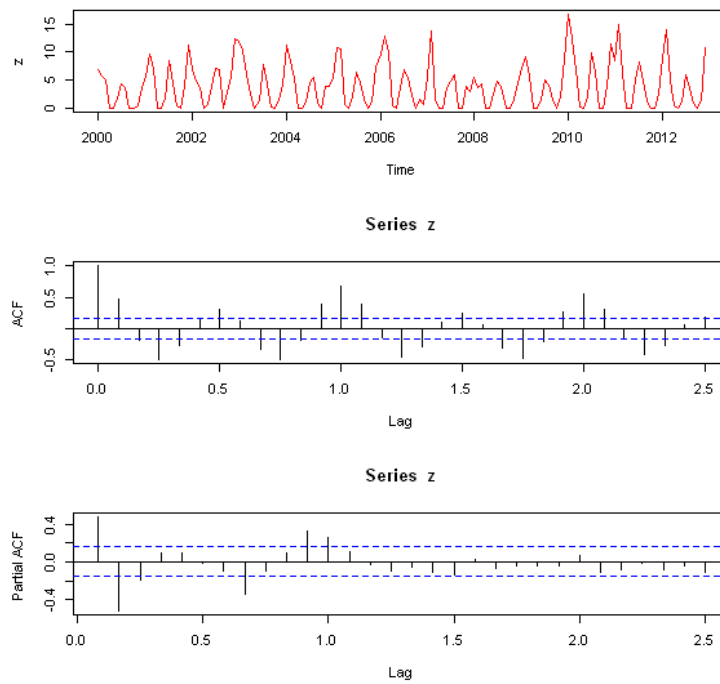
Joonis 24. Parandatud regressoritega ARIMAX mudeli parameetrite väärtused

4.1.3 Keerulise temperatuuri funktsiooni kasutamine regressor reana

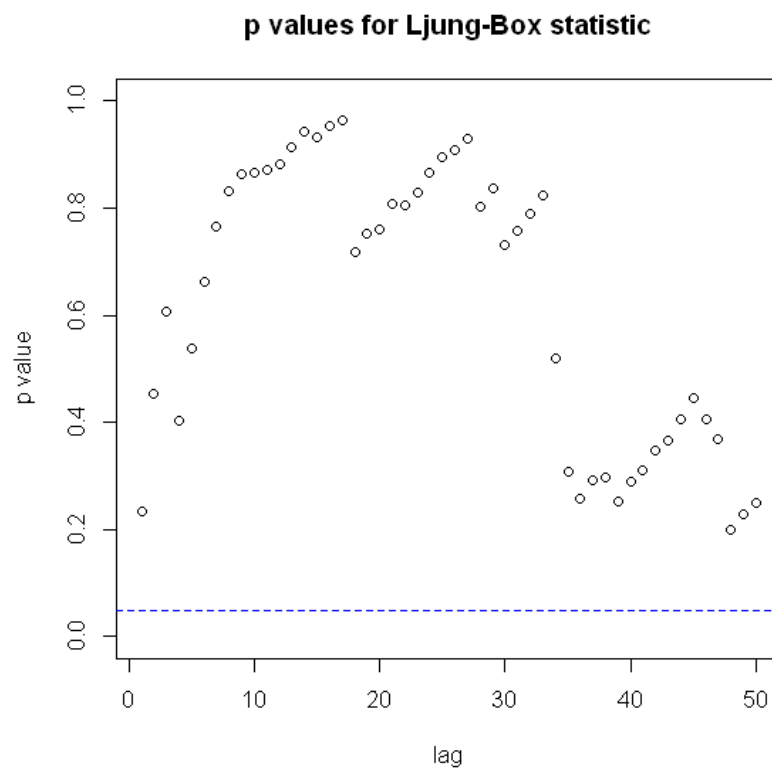
Kolmas temperatuuri funktsioon, mis on vaadeldavatest funktsioonidest kõige keerulisem on kujul

$$Temp_t = \begin{cases} T_t - 12, & \text{kui } T_t > 12 \\ 0, & \text{kui } 5 < T_t < 12 \\ 5 - T_t, & \text{kui } T_t < 5 \end{cases}.$$

Moodustame funktsioonist saadud temperatuuri väärtustest aegrea ning vaatleme vastava aegrea käitumist ajas. Saadud aegrida ei ole mõistlik enne regressorina kasutamist sesoonselt diferentsida samadel kaalutlustel nagu eelmiste funktsioonidest saadud aegride korral. Joonise 25 z graafikul on näha, et väärtused on positiivsed, mida me soovisimegi ning perioodile vastavad autokorrelatsioonid kahanevad. Vaadeldes osautokorrelatsioone (Joonis 25 *Partial ACF*), siis esimesed üksteist väärtust asuvad veapiiridest väljas. Seetõttu proovime andmetele sobitada AR(11) mudelit. Kui vastav mudel sobib andmetele, siis ei ole mõistlik vaadelda alternatiivseid mudeleid.



Joonis 25. Temperatuuri keerulise funktsioonist saadud väärtuste aegrea, ACF ja osa- ACF graafikud



Joonis 26. Mudeli AR(11) Ljung- Box p-väärtuste graafik

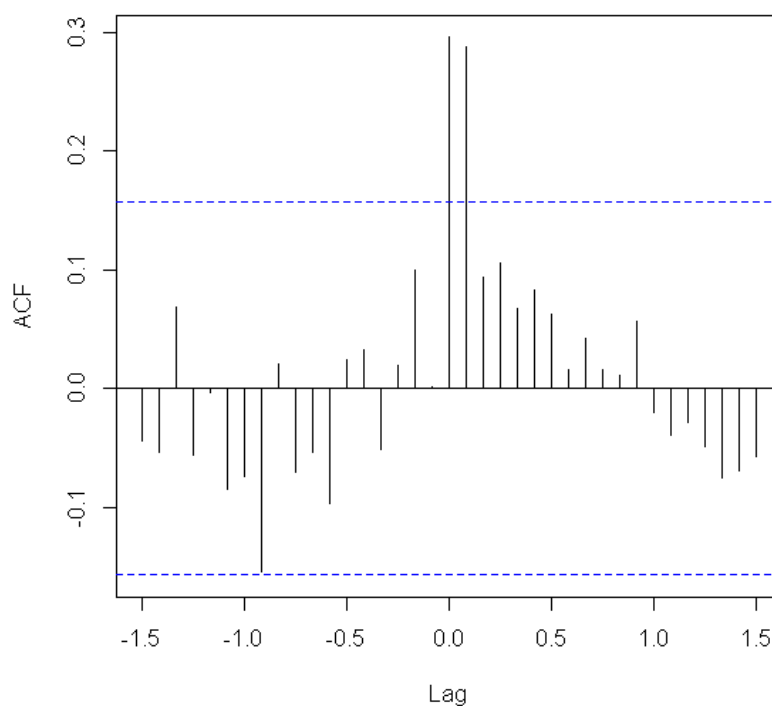
Tabel 9

Mudel	Ljung – Box väärtused	AIC väärtus	BIC väärtus
AR(11)	Kõik > 0.05	1195.96	1235.61

Ljung – Box p-väärtuste põhjal (Joonis 26) saab öelda, et mudel AR(11) sobib andmetele, kuna mudeli sobitamisel Ljung – Box teststatistiku väärtused on suuremad kui 0.05.

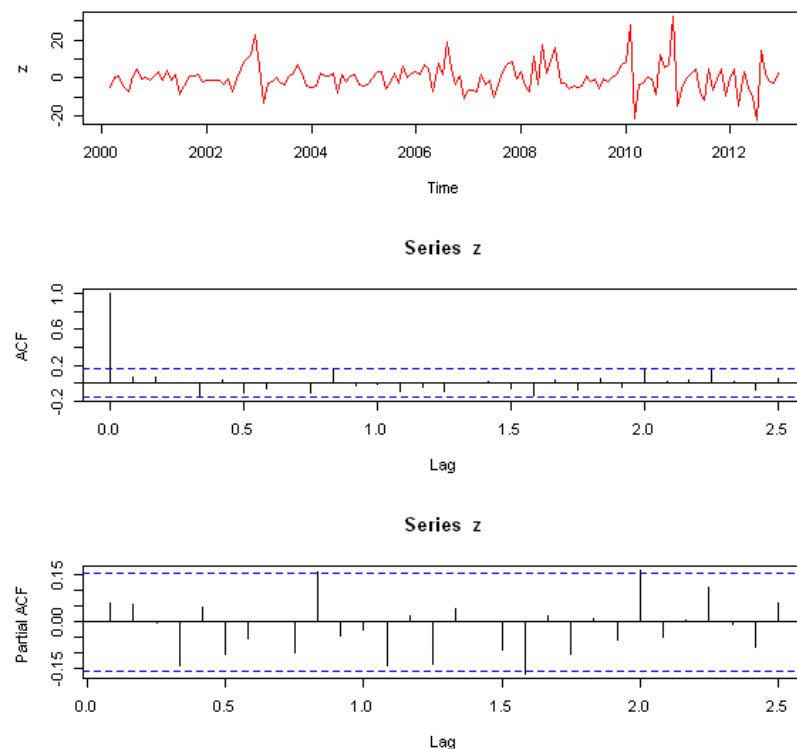
Lähtudes eelvalgedamise teisest sammust, siis sobitame AR(11) mudelit ka hindade aegreale, millelt on eelnevalt eemaldatud lineaarne trend. Peale sobitamist vaatleme mudelite jääkide vahelist ristkorrelatsioone (Joonis27), kus on näha, et kaks ristkorrelatsiooni on veapiiridest väljas – kohal null ja üks.

residuals(keeruline_mudel) & residuals(keertemp_mudel1)



Joonis 27. Vastavate mudelite jääkide vaheline ristkorrelatsioon

Püüame leida sobivat mudelit elektrihindade aegreale kasutades temperatuuri aegrida nihetega 0, 1 ning lisaks lineaarsele trendile vastavat regressorit. Sobitamise tulemusena tekkinud elektrihinna ennustamise jääkide autokorrelatsioonid ja osautokorrelatsioonid on toodud Joonisel 28. Vastava joonise *Partial ACF* graafikul on näha, et kõik osautokorrelatsioonide jäävad veapiiridesse, kuigi leidub kaks väärtust, mis on piiri peal. Joonise 28 ACF graafikul on samuti näha, et kõik autokorrelatsioonid jäävad veapiiridesse.



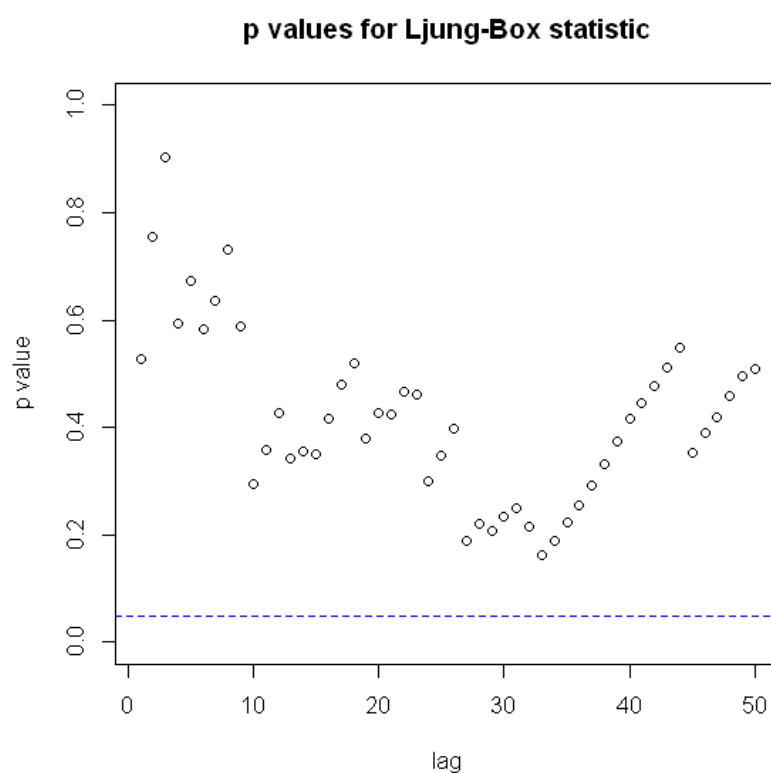
Joonis 28. Elektrihindade aegrida, ACF ja osa-ACF graafikud koos lineaarse trendi ja temperatuuri keerulise funktsiooni mõjudega

Joonise 28 lähtuvalt saame elektrihindade prognoosimisel aluseks võtta ajaloolised temperatuuri väärtustest moodustatud funktsiooni aegrida. Elektrienergia hinna aegreale sobitame seega mudelit kujul $ARIMA(0,0,0)$.

Tabel 10

Mudel	Ljung – Box väärtused	AIC väärtus	BIC väärtus
ARMA(0,0)	Kõik < 0.05	1072.57	1093.78

Esmane vaadeldav mudel ei sobinud, siis vaatleme alternatiivset mudelit, milleks võib olla AR(1), kuna üksikud osautokorrelatsiooni väärtused on piiri peal. Mudel AR(1) sobib andmetele Ljung – Box teststatistiku alusel (Joonis 29).



Joonis 29. AR(1) mudeli Ljung – Box teststatistiku väärtuste graafik

Uurime, kas mudel sobib kõik lisatud regressorid on olulised, selleks võrdleme liikmete kordajate väärtusi vastava väärtuse kahekordse standardhälbega.

Series: z
ARIMA(1,0,0) with non-zero mean

Coefficients:

	ar1	intercept	xreg	xreg1	xreg2
	0.7591	15.7132	0.5189	0.5758	0.1981
s.e.	0.0518	4.8313	0.1703	0.1672	0.0518

sigma² estimated as 54.51: log likelihood=-530.24
AIC=1072.48 AICc=1073.05 BIC=1090.74

Joonis 30. ARIMAX parameetrite väärtused

Joonisel 30 on näha, et kõik liikmed on olulised ja seega ei saa ühtegi liiget mudelist välja arvata.

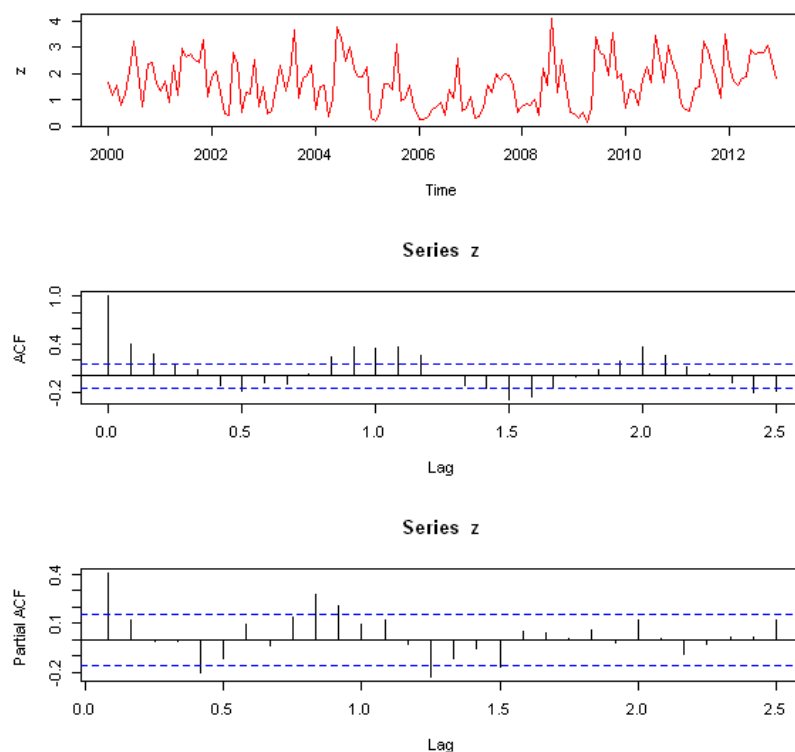
4.2 Sademete mõju Soome elektrienergia hindadele

Teiseks mõjuteguriks vaatleme sademete mõju Soome elektrienergia hindadele. Algsed andmed olid päevase intervalliga ja vastasid perioodile 2000 aastast kuni 2012 aasta lõpuni. Esmase töötluse käigus muutsime intervalli kuu andmeteks kasutades *Microsoft Excelis* käsku *average* (LISA 3). Analüüsimise etapid on samad, mis temperatuuri puhulgi, kuid erinevus on selles, et sademete andmetest me ei moodusta funktsioone. Funktsioonide asemel vaatlen kahte erijuhtu:

1. Talve kuude keskmiste sademete hulk asendada väärtusega null
2. Talve kuude keskmiste sademete hulga mõju avaldumine märtsi kuus, st summeeritakse kokku talvekuude sademete hulk ja lisatakse märtsikuule. Ülejäänud kuude väärtused on võrdsed nulliga.

Esmalt vaatleme originaalandmetest moodustatud aegrea käitumist (Jooni 31).

Joonisel 31 ACF graafikul on näha perioodilist käitumist, kuid vastavat aegrida ei ole mõistlik diferentsida sarnastel põhjustel nagu temperatuuri puhulgi. Joonisel 31 z graafikul ei ole silmaga näha trendi olemasolu, pigem paistab, et keskväärtsus on konstantne.



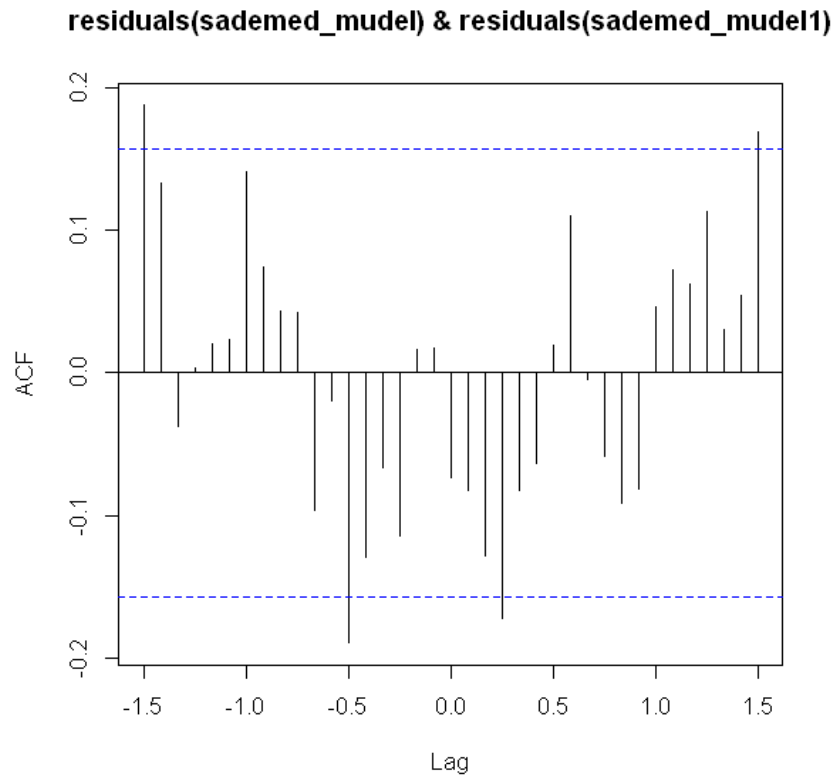
Joonis 31. Sademete väärtustest moodustatud aegrea, autokorrelatsiooni ja osautokorrelatsiooni graafik

Uurime kas sademete originaalandmetest moodustatud aegrida mõjutab Soome elektrienergia hinda. Kõigepealt püüame leida sademete aegreale sobiva mudeli. Joonise 31 osautokorrelatsioonide põhjal vaatleme AR(11) sobivust.

Tabel 11

Mudel	Ljung – Box väärtused	AIC väärtus	BIC väärtus
AR(11)	Kõik > 0.05	379.05	418.7

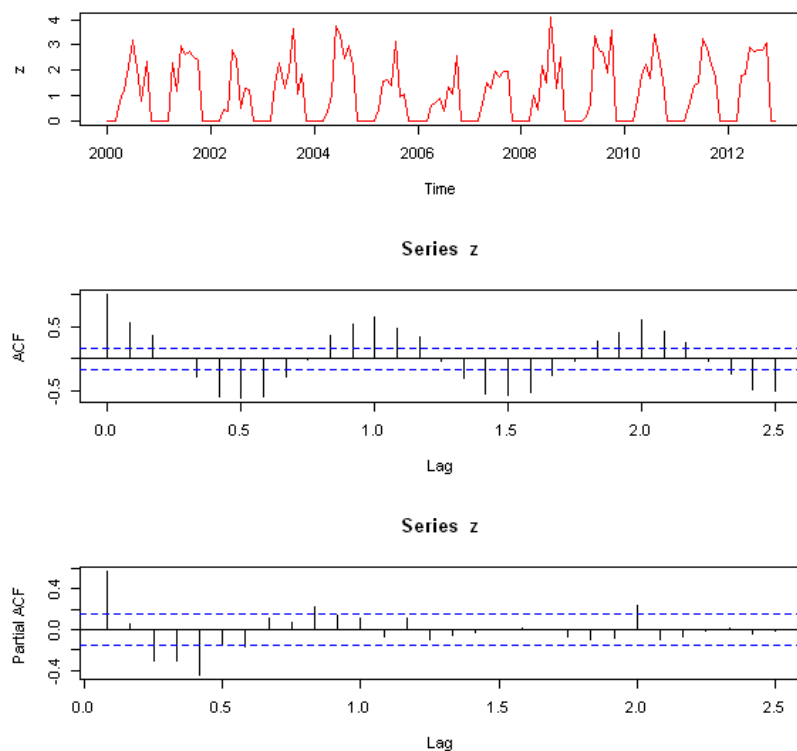
Tabelis 11 näeme, et vastav mudel sobib sademete aegreale. Seetõttu ei vaatle edasi alternatiivseid mudeleid. Sobitame AR(11) mudelit nüüd elektrihindade aegreale ja vaatleme vastavate mudeli jääkide ristkorrelatsioone.



Joonis 32. Sademete mudeli jääkide ja vastava elektrihinna mudeli jääkide vaheline ristkorrelatsioon

Joonise 32 on näha, et negatiivsel poolel on kaks ristkorrelatsiooni, mis on veapiiridest väljas, mille nihkeks on 15 ja 14. Suure tõenäosusega ei mõjuta pooleteise aasta tagused sademete hulk elektrienergia hinda. Positiivsel poolel on veel lisaks kolmas ristkorrelatsioon väljas. Ülekande funktsiooni eeldused ei ole täidetud ning sellest tulenevalt ei ole mõistlik sademete aegrida otseselt kasutada Soome elektrienergia hinna prognoosimisel.

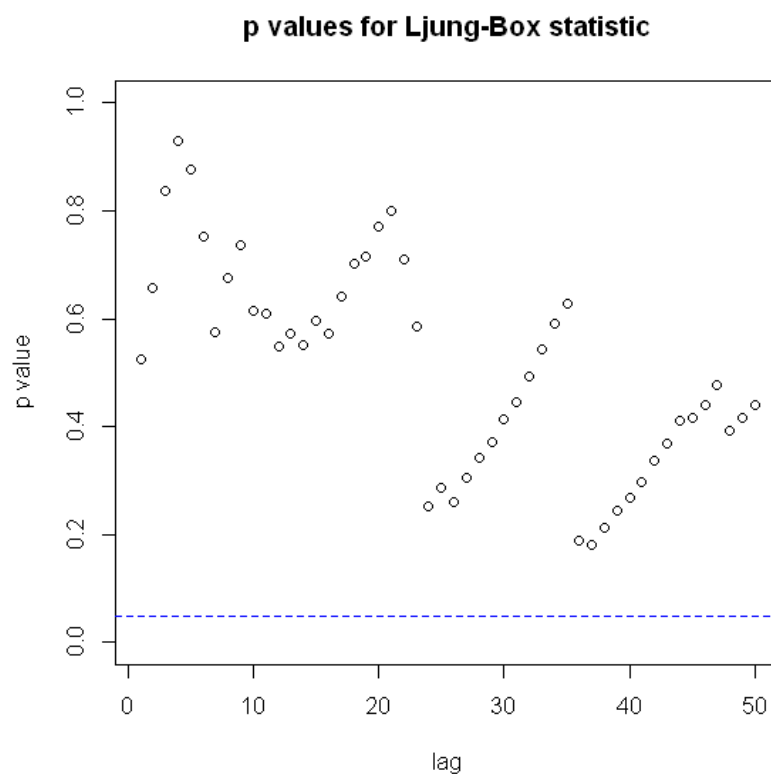
Vaatleme esimest olukorda, kus asendame iga aasta jaanuar kuni märts ja november kuni detsember sademete hulga väärtused nulliga.



1

Joonis 33. Korregeeritud sademete aegrea, ACF ja osa-ACF graafikud

Joonisel 33 osautokorrelatsiooni graafiku põhjal oleks mõistlik sobitada korregeeritud sademete andmetega aegrea alguses AR(5), kuna kuues ja seitsmes on piiri peal. Ljung – Box testi põhjal osutus mudel mitte sobivaks, seega lisame juurde liikmeid ja alternatiivseks mudeliks vaatleme AR(10). Ljung – Box testi põhjal võime lugeda mudeli sobivaks, kuna vastavad p-väärtused asuvad väljas pool veapiiri (Joonis 34).

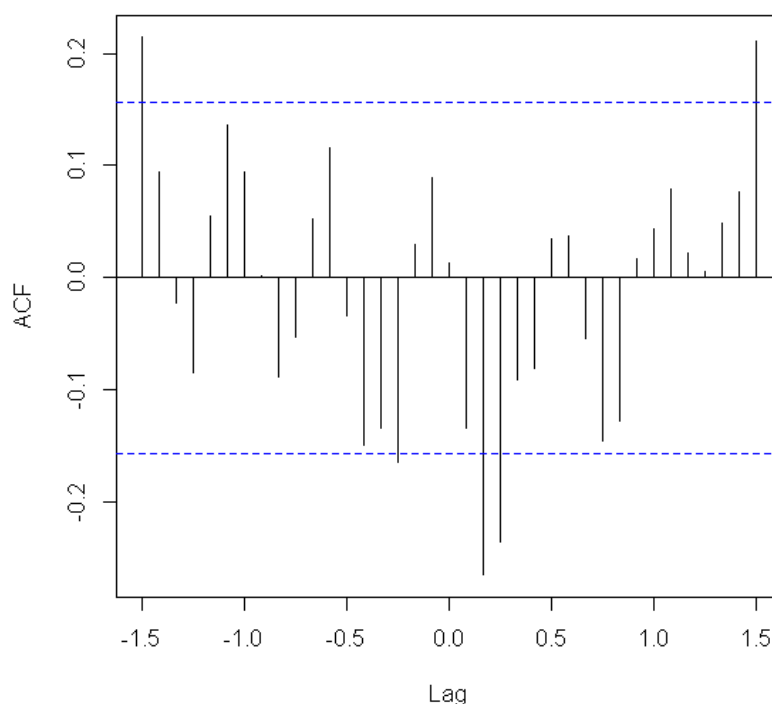


Joonis 34. Mudeli AR(11) Ljung – Box teststatistiku väärtuste graafik

Sobitades AR(11) mudelit elektriinna aegreale ning vaatleme vastavate mudelite vahelist ristkorrelatsioone (Joonis 35).

Sarnaselt eelmise juhuga paistab ka siin pooleteise aastaste mõju olemasolu olevat. Tegelikult arvesse võttes on selline mõju aga ebatõenäoline, et nii kauged mineviku väärtused mõjutavad elektrienergia hinda. Sellest tulenevalt nimetame pooleteise aasta taguseid ristkorrelatsioone juhuslikeks. Joonisel 35 on näha, et on teine ja kolmas ristkorrelatsioonid on selgelt veapiiridest väljas ning neid tasub kindlasti vaadelda kui sademete aegrea vastavate nihete mõju elektrienergia hinna aegreale.

residuals(sademed_mudel4) & residuals(sademed_mudel3)



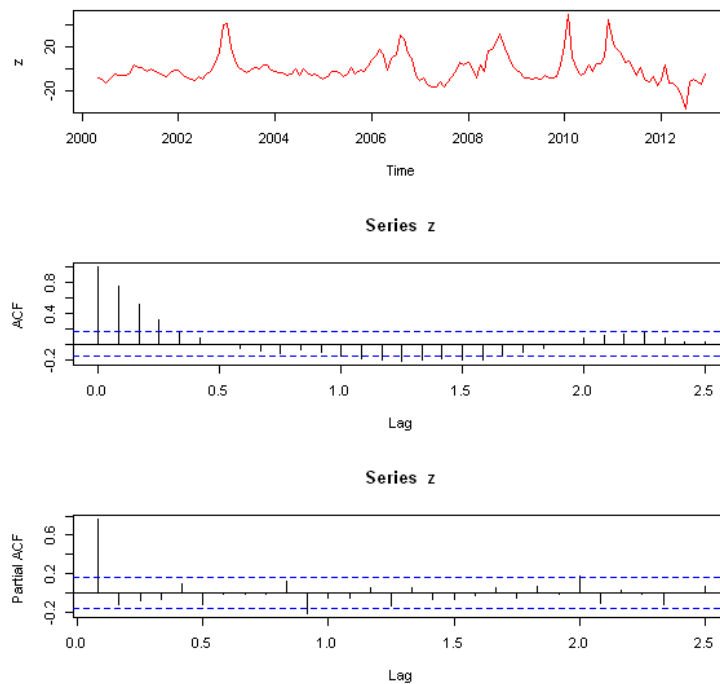
Joonis 35. Korreleeritud sademe aegrea mudeli ja vastava elektrihinna mudeli jääkide vaheline ristkorrelatsioon

Rakendades sademete sobitatud mudelit elektrienergia hindade aegreale, kasutades kolme regressorit, millest kaks on sademete aegrea nihked kaks ja kolm ning lineaarsele trendile vastav regressor.

Joonisel 36 on saadud regressorite mõjuga elektrienergia aegrida, vastavad autokorrelatsioonid ja osautokorrelatsioonid. Joonise 36 *Partial ACF* graafiku põhjal on mõistlik esmalt sobitada esimest järku autoregresiivset mudelit, kuna ainult üks osautokorrelatsioon on veapiirist väljas. Kui vastav mudel ei sobi, siis alternatiivseteks mudeliteks võib vaadelda segamudelit $ARMA(1,1)$, kuna autokorrelatsioonid kahanevad mingi arvuga korrutades. Mudeli $AR(1)$ sobivus on toodud Tabelis 12.

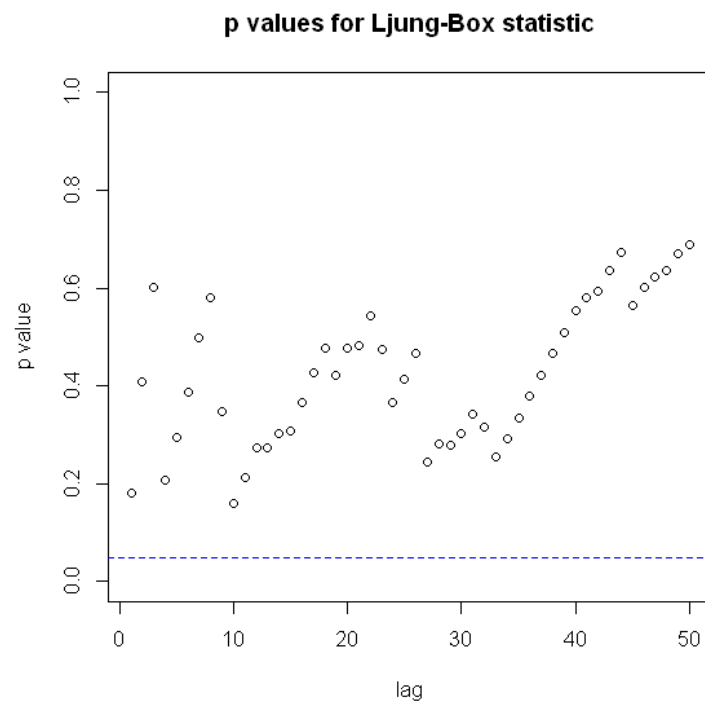
Tabel 12

Mudel	Ljung – Box väärtused	AIC väärtus	BIC väärtus
AR(1)	Kõik > 0.05	1076.94	1094.51



Joonis 36. Elektriinna mudel koos regressoritega

Tabelis 12 näeme, et vastav mudel sobib ning mudeli sobivust näeme ka Joonisel 37, kus Ljung – Box teststatistiku p-väärtused on veapiiridest väljas.



Joonis 37. AR(1) mudeli Ljung – Box teststatistikute p-väärtused

Ristkorrelatsioonide graafikul kolmas ristkorrelatsiooni väärtus on väiksem teisest, siis vaatleme, kas mõlemad nihked on mudeli jaoks olulised. Selleks võrdleme regressorite kordajate väärtusi vastavate väärtuste standardhälbetega (Joonis 38).

```
Series: z
ARIMA(1,0,0) with non-zero mean

Coefficients:
      ar1 intercept  xreg1  xreg2  xreg3
      0.7636  20.2190 -0.0786  0.8777  0.1929
s.e.  0.0515   5.4234  0.6940  0.7407  0.0589

sigma^2 estimated as 63.99: log likelihood=-532.18
AIC=1076.37  AICc=1076.94  BIC=1094.51
```

Joonis 38. ARIMAX mudeli parameetrite kordajate väärtused

Joonisel 38 näeme, et ebaolulisteks on xreg1 ja xreg2, seega võime need mudelist välja jätta. Tulemuseks saime, et sademete aegrea nihked ei ole olulised, vaid oluliseks osutus lineaarse trendile vastav regressor. Joonisel 39 on ARIMAX mudeli oluliste liikmete parameetrite väärtused.

```
Series: z
ARIMA(1,0,0) with non-zero mean

Coefficients:
      ar1 intercept  xreg3
      0.7594  21.4891  0.1940
s.e.  0.0519   5.1820  0.0581

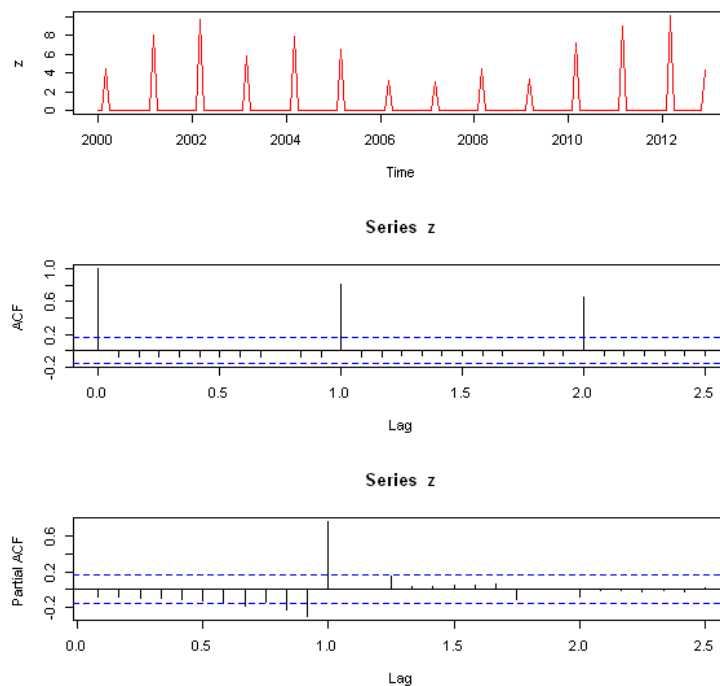
sigma^2 estimated as 64.68: log likelihood=-532.99
AIC=1073.98  AICc=1074.25  BIC=1086.07
```

Joonis 39. Lineaarsele trendile vastava regressoriga hinna mudeli parameetrite väärtused

Teisel juhul vaatleme olukorda, kus talvekuude sademete hulk mõjutaks elektri hindu märtsikuus. See on loogiline järeldus, kuna suurvee ajal peaks hüdroenergia mõju olema suurem.

Moodustame teisest korrigeeritud sademete andmestikust aegrea ja vaatleme vastavate autokorrelatsioonide ja osautokorrelatsioonide käitumist. Joonisel 40 näeme, et autokorrelatsioonide graafikul esineb ainult perioodile vastavas osas

esimese autokorrelatsiooni veapiirist välja jäämise. *Partial Acf* graafikult näeme, aga, et kuni kuuendani osautokorrelatsioonini püsivad väärtused veapiiride sees, kuid seejärel toimub väike piirist väljajäämised. Perioodile vastavas osas on ainult üks osautokorrelatsioon, mis jääb veapiirist välja.



Joonis 40. Teine korrigeeritud sademete andmestiku aegrida, ACF ja osa-ACF

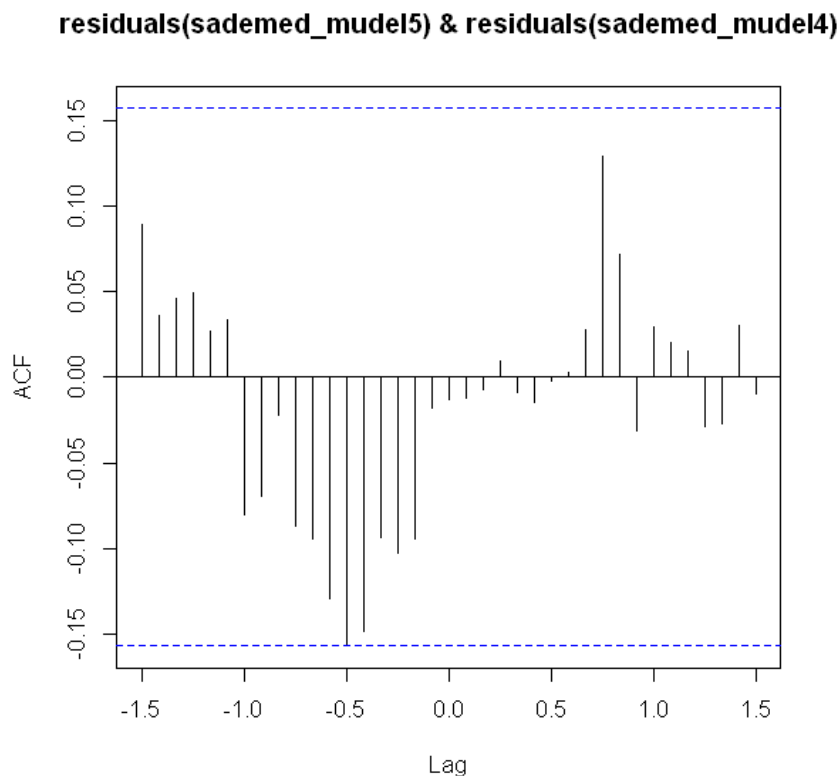
Kokkuvõttes oleks mõistlik vastavale aegreale sobitada ainult perioodile vastavat mudeli osa, st mudelit kujul $ARIMA(0,0,0)x(1,0,0)$.

Tabel 13

Mudel	Ljung – Box väärtused	AIC väärtus	BIC väärtus
ARIMA(0,0,0)x(1,0,0)	Kõik > 0.05	1369.3	1378.45

Tabelis 13 mudel sobib korrigeeritud sademete andmetele, kuna Ljung _ Box p-väärtused on suuremad kui 0.05. Sobitame vastavat mudelit hindade aegreale, millelt on eemaldatud enne lineaarne trend. Vaatleme vastavate mudeli jääkide ristkorrelatsioone (Joonis 41), kus on näha, et kõik ristkorrelatsioonid asuvad veapiiride

sees. Seega ei sobi korrigeeritud sademete andmetest saadud aegrida Soome elektrienergia hinna prognoosimiseks.



Joonis 41. Ristkorrelatsioonid vastavate mudelite jääkide vahel

4.3 ARIMAX mudeli väljakirjutamine

Eelmistes punktides vaatlesime temperatuuri ja sademete mõju Soome elektrienergia hindadele, kasutades temperatuuri puhul kolme erinevat funktsiooni ning sademete puhul korrigeerisime andmestiku kahel erineval viisil. ARIMAX tüüpi mudelite sobitamise teel saime mitu erinevat mudelit. Enne sobiva mudeli väljakirjutamist võrdleme vastavaid mudeleid omavahel. Võrdlemiseks kasutame tuntuid mõõdikuid – keskmine absoluutne viga (MAD), keskmine ruutviga (MSE), ruutkeskmine viga (RSME) ja keskmine suhteline viga (MAPE). Samuti kasutame võrdlemiseks Akaike ja Bayesi Informatsiooni kriteeriumit (BIC), kuid viimast on kasutatav juhul, kui mudelid on sobitatud samadele andmetele.

Lihtsuse mõttes tähistame temperatuuri funktsiooni $(T_t - \mu)^2$ sobivat mudelit $ARX(1)_{ruut}$, $|T_t - \mu|$ funktsioonile vastavat mudelit kujul $ARX(1)_{abs}$ ja keerulisema funktsiooniga kirjeldatud temperatuuri andmete mudelit $ARX(1)_{keeruline}$ ning esimesel viisil korrigeeritud sademetega vastavat mudelit $AR(1)_{sademed}$.

Mudelite võrdlus ja vastavate mõõdikute väärtused on näidatud Tabeli 14.

Tabel 14

Mõõdikud	Mudel	$ARX(1)_{ruut}$	$ARX(1)_{abs}$	$ARX(1)_{keeruline}$	$AR(1)_{sademed}$
MAD		5.081535	5.254881	5.029219	5.42539
MSE		54.38297	58.87876	54.5119	64.68072
RSME		7.374481	7.67325	7.383218	8.042432
MAPE		0.1452204	0.1474591	0.143524	0.1479781
AIC		1065.31	1070.61	1072.48	1073.98
BIC		1083.53	1088.79	1090.74	1086.07

Tabeli 14 põhjal parimat mudelit välja tuua on raske, kuna ei ole ühtegi mudelit, mille kõik mõõdikud oleksid teistest paremad. Parimaks ARIMAX tüüpi mudeliks loeme mudelit, millel on rohkem mõõdikuid, mis oleks võrreldes teistega väiksemad. Seega parimaks sobivaks mudeliks saime $ARX(1)_{ruut}$ (Joonis 19). Mudeli väljakirjutamisel kasutame definitsiooni 7 põhjal saadud valemit kujul

$$Z_t = \sum_{i=1}^r \delta_i Z_{t-i} + \sum_{i=0}^s v_i X_{t-b-i} + \varepsilon_t.$$

ARIMAX mudeliks saame:

$$(1 - 0.7657B)(Y_t - 0.0314X_{t-1} - 0.0343X_t) = A_t.$$

ehk

$$(1 - 0.7657B)Y_t - 0.0314X_{t-1} + 0.0240X_{t-2} - 0.0343X_t + 0.0263X_{t-1} = A_t.$$

ehk

$$Y_t = 0.7657Y_{t-1} + 0.0314X_{t-1} - 0.0240X_{t-2} + 0.0343X_t - 0.0263X_{t-1} = A_t$$

Asendame $Y_t = Z_t - c_1t - c_0$ saame

$$\begin{aligned} Z_t - 0.1914t - 16.4402 \\ = 0.7657(Z_{t-1} - 0.1914(t-1) - 16.4402) + 0.0343X_t \\ + 0.0051X_{t-1} - 0.0240X_{t-2} + A_t \end{aligned}$$

Peale lihtsustamist saame

$$\begin{aligned} Z_t = 0.7657Z_{t-1} + 0.0343X_t + 0.0051X_{t-1} - 0.0240X_{t-2} + 0.0448t + 3.9985 \\ + A_t \end{aligned}$$

4.4 Sobiva ARIMAX mudeliga lühiajaline ennustamine

Eelmises alapeatükis saime ARIMAX mudeli, mis oli saadud Soome elektrienergia hinna aegrea mudeldamisel kasutatud lihtsa temperatuuri funktsiooni aegrea nihkeid ja lineaarsele trendile vastavat regressorit. Kasutades sobitatutest parimat ARIMAX mudelit saame leida Soome elektrienergia hindade prognoose, mis on Tabelis 15.

Tabel 15

Aeg	Lo 95	Lo 80	Väärtus	Hi80	Hi95
Jaanuar 2013	30.71557	36.07510	46.16362	56.25215	61.61168
Veebruar 2013	18.65686	25.68777	38.92242	52.15707	59.18798

Tänase seisuga on meil olemas prognoositud tegelikud väärtused. Mudel kasutas ajaloolisi andmeid kuni 2012 aasta detsembrini.

Võrdleme prognoositud väärtusi tegelike väärtustega, siis Tabelis 16 näeme, kui palju prognoositud väärtused erinevad reaalistest väärtustest. Erinevus jaanuarikuus on suurem kui veebruaris, kuid võib öelda, et saadud ARIMAX mudel prognoosib küllaltki

hästi lähituleviku väärtusi, kasutades temperatuuri väärtustest moodustatud funktsiooni aegreale vastavaid regressoreid.

Tabel 16

Aeg	Prognoositud väärtus	Tegelik väärtus	Erinevus
Jaanuar 2013	46,16362	41,59	-4,57362
Veebruar 2013	38,92242	39,43	0,50758

Kokkuvõte

Käesoleva töö eesmärgiks oli otsida võimalikult lihtsaid, kuid siiski teatud mõttes optimaalseid mudeleid elektrienergia hinna lühiajaliseks prognoosimiseks. Kuna Eesti kohta ei ole veel piisavalt andmeid, viidi vastav analüüs läbi Soome kohta käivate andmestike baasil.

Kõigepealt otsisime mudeleid, mis kasutaks ainult elektrienergia hinna ajaloolisi andmeid. Mõnevõrra üllatavalt selgus, et andmetes ei ole täheldada märgatavat sesoonset sõltuvust ning parimaks mudeliks osutus lineaarse trendiga lihtne AR(1) tüüpi mudel. Edasi uurisime prognooside täpsustamise võimalusi juhul, kui on olemas andmestikud elektrienergia hinna kujunemise seisukohalt asjakohaste väliste tegurite, nimelt temperatuuri ja sademete kohta. Selgus, et temperatuuril on selge mõju energiahinna kujunemisele olemas ning adekvaatse info olemasolul tulevaste kuude keskmise temperatuuri kohta on võimalik elektrienergia hinda paremini prognoosida. Tulevase kuu hinda on võimalik veidi paremini prognoosida ka siis, kui me kasutame regressorina jooksva kuu keskmist temperatuuri.

Teise asjakohase välisteguriks uuriti sademete mõju hinnale, kuid selgus, et otsest mõju täheldada ei saa. Kuna sademete hulk hüdroenergiale on väike ning suuremalt võttes on hüdroenergia osakaal kogu energiatootmisest minimaalne, siis sellest tulenevalt on saadud järeldus ka loomulik.

Käesoleavas töös võeti arvesse ainult mõningaid faktoreid, mis võivad elektrienergia hinnale mõju avaldada. On olemas palju keerukamaid ökonomeetrilisi mudeleid, kuid kahjuks ei olnud selle töö autoritel võimalik oma tulemusi keerukamate mudelite prognoosidega võrrelda. Samas on saadud tulemused kindlasti kasutatavad selleks, et täielikumate mudelite loomisel saaks praktikud ja teoreetikud kontrollida, kas mudeli suurem keerukus ikkagi prognooside parandamise mõttes ennast ära tasub.

Models and short time forecasts for electricity prices in Finland

Anni Niidumaa

Summary

Modeling electricity prices is of great interest in all countries that have joined or are planning to join with an electric power market. Estonia has also joined a power market system, various companies related to producing, selling and distributing electric power, are developing their own systems for predicting future prices. There are various models developed for the purpose of describing the evolution of the price of electricity and for providing forecasts for different time spans. Some of the models are extremely complicated and try to take into account all major factors that may influence the price of the electricity but it is not very clear if such comprehensive models are better for practical forecasting than simple models.

The aim of this thesis is to develop some simple models for predicting future electricity prices in order to provide a basis of comparison for developers and implementers of more complicated models. This is done on the basis of electricity prices in Finland since there is not enough data about the market prices in Estonia yet.

The thesis is divided into 4 chapters. The first chapter is about some background information of the electricity market, different factors affecting electricity prices, the history and development of the electricity market in Estonia. In Chapter 2, are introduced the main time series models that are used in this thesis. In Chapter 3, is investigated the problem of prediction the future electricity prices by using only information about historical prices. The best models from the class of ARIMA models is founded both for the original and for transformed (using the logarithmic

transformations) data. These two approaches are compared and the best approach is also determined. Surprisingly it turned out that there is no clear evidence of seasonality in the data and that a simple model with linear trend with AR(1) errors was the best one.

In the final chapter the possibility of using exogenous regressors of temperature and precipitation time series for improving the predictions. It is determined that it is possible to improve predictions if good predictions for average temperatures of future months are available and that by knowing the average temperature of the last month it is possible to improve the prediction of the price of electricity for the next months.

Kasutatud kirjandus

- [1] Eleringi, 2012. Eesti elektrituru täielik avanemine nr2/2012(4)
- [2] Elering, 2012. Eesti Elektrisüsteemi varustuskindluse aruanne 2012
- [3] Elering, Tallinn 2012. Elektrituru käsiraamat
- [4] http://ecocrete.eu/ET_24_4.htm
- [5] <https://www.energia.ee/et/polevkivi>
- [6] <http://www.envir.ee/1188363>
- [7] <http://www.esha.be/energy/facts.html>
- [8] http://en.wikipedia.org/wiki/Akaike_information_criterion#Definition
- [9] http://en.wikipedia.org/wiki/Bayesian_information_criterion
- [10] <http://issuu.com/maeinstituut/docs/polevkivi?mode=window&pageNumber=1>
- [11] <http://www.e24.ee/940550/eesti-energia-pole-toetust-pole-taastuenergiat/>
- [12] <http://faculty.washington.edu/ezivot/econ584/notes/unitroot.pdf>
- [13] http://issuu.com/elering/docs/aruanne_elektri_ja_gaasiturust_eestis?mode=window&pageNumber=1 (lk 8)
- [14] Massimiliano Serati, Matteo Manera, Michele Plotegher, 2008. Modeling electricity prices: From the State of the Art to a Draft of a New Proposal
- [15] <http://www.mathworks.se/help/econ/autocorrelation-and-partial-autocorrelation.html>
- [16] https://moodle.ut.ee/pluginfile.php/153537/mod_resource/content/21/aegread.pdf
- [17] <https://onlinecourses.science.psu.edu/stat510/?q=node/75>

- [18] <http://www.tuumaenergia.ee/index.php?id=60>
- [19] <http://www.tuumaenergia.ee/index.php?id=107>
- [20] <http://www.world-nuclear.org/info/Economic-Aspects/Economics-of-Nuclear-Power/#.UWPUDjc6VII>

Appendixes

Lisa 1. Soome elektrienergia hinnad

year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2013	41,59	39,43	45,01	-	-	-	-	-	-	-	-	-
2012	38,82	52,81	36,51	36,48	33,34	27,38	13,67	38,18	41,03	38,57	36,95	46,79
2011	68,92	64,58	60,91	52,93	54,42	48,54	42,20	48,98	38,86	36,90	42,03	33,34
2010	65,78	93,70	55,22	43,71	39,47	41,96	48,76	43,21	51,20	51,23	56,63	91,34
2009	41,08	38,33	34,88	34,46	33,13	35,38	33,81	37,27	35,60	35,08	36,69	47,98
2008	46,12	39,74	31,94	43,56	38,37	57,62	59,07	65,25	73,37	60,35	52,45	44,36
2007	27,46	30,08	23,70	22,19	22,02	26,92	22,37	26,88	32,24	37,24	45,59	43,63
2006	43,09	47,75	53,24	49,14	35,17	45,33	49,14	67,24	63,72	50,78	46,43	32,02
2005	23,19	25,59	30,76	30,89	30,64	26,74	29,99	35,07	29,97	33,86	32,07	37,00
2004	28,53	27,03	27,28	25,81	26,74	31,10	25,93	31,67	28,08	26,37	28,13	25,48
2003	69,84	47,01	37,57	29,61	28,06	26,16	28,14	31,25	29,93	33,70	33,87	28,93
2002	24,91	20,41	18,62	17,39	15,85	19,93	18,39	22,76	25,81	31,54	43,25	67,68
2001	20,46	27,13	25,78	26,10	21,77	25,16	22,32	21,14	19,44	18,87	21,59	24,72
2000	18,69	12,92	12,49	13,01	14,31	12,68	9,77	14,61	18,70	16,73	17,21	17,34

LISA 2. Logaritmitud Soome elektrienergia hinnad

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2013	3.727860	3.674527	3.806885									
2012	3.658936	3.966701	3.597586	3.596764	3.506758	3.309813	2.615204	3.642312	3.714304	3.652475	3.609566	3.845670
2011	4.232946	4.167905	4.109397	3.968970	3.996732	3.882388	3.742420	3.891412	3.659965	3.608212	3.738384	3.506758
2010	4.186316	4.540098	4.011325	3.777577	3.675541	3.736717	3.886910	3.766072	3.935740	3.936325	4.036539	4.514589
2009	3.715521	3.646233	3.551914	3.539799	3.500439	3.566147	3.520757	3.618189	3.72346	3.557631	3.602504	3.870784
2008	3.831247	3.682358	3.463859	3.774139	3.647276	4.053870	4.078723	4.178226	4.295515	4.100161	3.959860	3.792338
2007	3.312730	3.403860	3.165475	3.099642	3.091951	3.292870	3.107721	3.291383	3.473208	3.617383	3.819688	3.775745
2006	3.763291	3.865979	3.974810	3.894673	3.560193	3.813969	3.894673	4.208268	4.154498	3.927503	3.837946	3.466361
2005	3.143721	3.242202	3.426215	3.430433	3.422306	3.286161	3.400864	3.557346	3.400197	3.522234	3.467921	3.610918
2004	3.350956	3.296947	3.306154	3.250762	3.286161	3.437208	3.255401	3.455370	3.335058	3.272227	3.336837	3.237894
2003	4.246207	3.850360	3.626206	3.388112	3.334345	3.264232	3.337192	3.442019	3.398861	3.517498	3.522530	3.364879
2002	3.215269	3.016025	2.924236	2.855895	2.763170	2.992226	2.911807	3.125005	3.250762	3.451257	3.766997	4.214791
2001	3.018472	3.300640	3.249599	3.261935	3.080533	3.225255	3.105483	3.051167	2.967333	2.937573	3.072230	3.207613
2000	2.927989	2.558776	2.524928	2.565718	2.660959	2.540026	2.279316	2.681706	2.928524	2.817204	2.845491	2.853016

LISA 3. Keskmiste temperatuuri ja sademete tabel

aasta	kuu	keskmise temperatuur	keskmise sademete hulk
2000	1	-1,929677419	1,638387097
2000	2	-0,777586207	1,183275862
2000	3	-0,04483871	1,554193548
2000	4	7,764333333	0,782166667
2000	5	10,425	1,249032258
2000	6	13,8825	2,118666667
2000	7	16,33193548	3,212741935
2000	8	15,67145161	2,100483871
2000	9	10,24683333	0,757333333
2000	10	9,137580645	2,341451613
2000	11	4,569	2,429
2000	12	1,341774194	1,608548387
2001	1	-0,844032258	1,320645161
2001	2	-4,613392857	1,712857143
2001	3	-2,176290323	0,916935484
2001	4	5,9075	2,284333333
2001	5	9,836129032	1,173548387
2001	6	13,86416667	2,943166667
2001	7	20,42145161	2,606612903
2001	8	16,5933871	2,733548387
2001	9	12,45166667	2,5105
2001	10	8,773709677	2,432096774
2001	11	1,0805	3,2875
2001	12	-6,205806452	1,12983871
2002	1	-1,799677419	1,863870968
2002	2	0,173571429	2,109464286
2002	3	1,351774194	1,255806452
2002	4	5,557333333	0,471166667

2002	5	12,60435484	0,403870968
2002	6	15,885	2,807333333
2002	7	19,00887097	2,389677419
2002	8	18,84048387	0,517258065
2002	9	11,8655	1,288666667
2002	10	2,080322581	1,231451613
2002	11	-0,439833333	2,522666667
2002	12	-7,453064516	0,758870968
2003	1	-6,846935484	1,498548387
2003	2	-5,414642857	0,460714286
2003	3	-1,04	0,578709677
2003	4	2,568833333	1,448333333
2003	5	10,42080645	2,324354839
2003	6	13,28166667	1,3145
2003	7	19,90419355	2,006451613
2003	8	16,26354839	3,644032258
2003	9	12,33245402	1,0775
2003	10	4,897408602	1,832258065
2003	11	3,323149425	1,9505
2003	12	0,837903226	2,291451613
2004	1	-6,281129032	0,608225806
2004	2	-3,545689655	1,485689655
2004	3	-0,285967742	1,557903226
2004	4	5,267166667	0,3505
2004	5	10,11370968	1,062419355
2004	6	13,15143103	3,783166667
2004	7	16,66725806	3,352741935
2004	8	17,44080645	2,457580645
2004	9	12,9765	2,988
2004	10	6,874354839	2,175
2004	11	1,013166667	1,896666667
2004	12	1,047741935	1,854677419
2005	1	-0,567741935	2,252258065

2005	2	-5,748214286	0,316642857
2005	3	-5,584193548	0,187634409
2005	4	4,363683908	0,456316092
2005	5	10,17709677	1,582419355
2005	6	14,11716667	1,609166667
2005	7	18,52435484	1,385322581
2005	8	16,66564516	3,115806452
2005	9	13,49483333	0,975333333
2005	10	7,937258065	1,075875576
2005	11	3,781833333	1,5755
2005	12	-2,409032258	0,750806452
2006	1	-4,614516129	0,22884993
2006	2	-7,783035714	0,258571429
2006	3	-4,819032258	0,355645161
2006	4	4,619666667	0,6175
2006	5	10,35645161	0,722741935
2006	6	15,7075	0,911166667
2006	7	18,93322581	0,402580645
2006	8	17,51790323	1,367903226
2006	9	14,244	1,0475
2006	10	9,137580645	2,585806452
2006	11	3,2975	0,602833333
2006	12	4,288225806	0,688387097
2007	1	-0,589193548	1,125645161
2007	2	-8,692678571	0,299464286
2007	3	3,382903226	0,375
2007	4	4,901166667	0,684666667
2007	5	11,23096774	1,53516129
2007	6	15,72483333	1,296
2007	7	16,87870968	1,988064516
2007	8	18,01532258	1,752741935
2007	9	11,86566667	1,9675
2007	10	7,38516129	1,948258065

2007	11	1,195666667	1,569968648
2007	12	1,945	0,546612903
2008	1	-0,410413978	0,742096774
2008	2	1,266034483	0,833275862
2008	3	0,716206897	0,776774194
2008	4	6,417609195	1,009148148
2008	5	10,24629032	0,430806452
2008	6	14,40583333	2,211637931
2008	7	16,74032258	1,539032258
2008	8	15,99354839	4,072903226
2008	9	10,56010714	1,28
2008	10	9,028548387	2,54516129
2008	11	3,344	1,5385
2008	12	0,394193548	0,536290323
2009	1	-2,084510753	0,479091398
2009	2	-4,174107143	0,310714286
2009	3	-0,795967742	0,518548387
2009	4	5,319	0,148405797
2009	5	10,77419355	0,688037634
2009	6	13,61033333	3,371166667
2009	7	17,02032258	2,817399889
2009	8	16,20483871	2,701612903
2009	9	13,647	1,904074713
2009	10	5,101612903	3,57233871
2009	11	2,933	1,79525
2009	12	-3,678709677	1,998870968
2010	1	-11,66435484	0,712580645
2010	2	-7,485178571	1,383214286
2010	3	-1,797258065	1,315207373
2010	4	4,8605	0,816017241
2010	5	11,49258065	1,777413978
2010	6	14,13033333	2,233666667
2010	7	21,84032258	1,685290323

2010	8	18,33612903	3,418629032
2010	9	11,74883333	2,588007902
2010	10	5,188064516	1,651774194
2010	11	1,005666667	3,088850575
2010	12	-6,580645161	2,437758065
2011	1	-3,55983871	1,979124027
2011	2	-9,961071429	0,855446429
2011	3	-1,258225806	0,610269744
2011	4	5,187666667	0,570333333
2011	5	10,27403226	1,427241935
2011	6	17,05666667	1,52125
2011	7	20,2616129	3,25233871
2011	8	16,91580645	2,873306452
2011	9	13,4572619	2,231333333
2011	10	8,309516129	1,746290323
2011	11	4,934333333	1,043083333
2011	12	2,469677419	3,471532258
2012	1	-3,952258065	2,329516129
2012	2	-8,899655172	1,69355364
2012	3	-0,000645161	1,533145161
2012	4	4,466833333	1,817416667
2012	5	11,02193548	1,851048387
2012	6	13,1505	2,8895
2012	7	17,90094086	2,754193548
2012	8	15,60612903	2,782822581
2012	9	12,83316667	2,801166667
2012	10	6,712258065	3,063548387
2012	11	3,577666667	2,48325
2012	12	-5,871698925	1,807096774

LISA 4. Kasutatud skript R's

```
graafikud=function(z,mitu=30){
  layout(1:3) #jaotab graafikaekraani kolmeks võrdseks osaks
  plot(z, col=2) #esimesesse ossa aegrea graafik
  acf(z,mitu) #teiseks autokorrelatsioonid kuni etteantud nihete arvuni
  pacf(z,mitu) #osaaautokorrelatsioonid
  layout(1) #graafikaaken edasisteks käskudeks jälle ühes osas
}

korrelatsioonid=function(z,mitu=30){
  layout(1:2) #jaotab graafikaekraani kolmeks võrdseks osaks
  acf(z,mitu) #teiseks autokorrelatsioonid kuni etteantud nihete arvuni
  pacf(z,mitu) #osaaautokorrelatsioonid
  layout(1) #graafikaaken edasisteks käskudeks jälle ühes osas
}

library(forecast)

#####

andmed=read.table("f:/UT/loputoo/Elspot prices.csv",sep=";",na.strings="-",header=F,skip=1,dec=",")

head(andmed)

#andmed tabel kujul ->üheks tulbaks

rida=c(t(andmed[13:2])) # pärast ajasümber pööramist oleksid andmed ajalises järjekorras

rida

# pöörame ajas ümber ja aegrea moodustamine

rida1=rida[length(rida):1]

#puuduvate andmete eemaldamine

rida=rida1[!is.na(rida1)]

#aegrida

hind=ts(rida,start=2000,frequency=12)

#####
```

```

hind

plot(hind, col=2,xlab="Aeg",ylab="Elektrienergia hind EUR/MWh")

#####

#trend

n=length(hind)

aeg=1:n

lin_trend=lm(hind~aeg)

#kasutame arima

lin_trend=arima(hind,xreg=aeg)

lin_trend

#trendiks on : 20.7958+0.1915*t; teine rida annab meile ebatäpsuse

jääk=residuals(lin_trend)

#jääb järgi kui indeks maha võtta: jäägid=rida-trend => trend=rida-jäägid

trend1=hind-jääk

lines(trend1,col="4")

plot(jääk, col=2,xlab="aeg",ylab="Elektrienergia hind eur/MWh pärast lineaarse trendi eemaldamist")

#####

#mudeli leidmine

korrelatsioonid(hind,50)

korrelatsioonid(jääk,50)

#acf kerge sesoonsus sees, kuid acf lähevad keskmise kiirusega 0

#osa acf esimene on selgelt veapiiridest väljas

#perioodile vastav osas on acf esimene( 2 juures) veapiiridest väljas, teised veapiiride piirides

# alguses proovime otse sobitada mudelit, st madalamat järku mudelit

PP.test(hind)

#diferentsitud rida ei ole parem kui algne rida

m0=arima(hind,xreg=aeg,order=c(1,0,0))

tsdiag(m0,50)

```

```

#Ljung-Boxi kõik väärtused veapiiridest väljas -> mudel sobib hästi

m1=arima(hind,xreg=aeg,order=c(0,0,4))

tsdiag(m1,50)

m2=arima(hind,xreg=aeg,order=c(1,0,1))

tsdiag(m2,50)

m3=arima(hind,xreg=aeg,order=c(0,0,4),seasonal=c(0,0,2))

tsdiag(m3,50)

m0 #parim
m1
m2
mean(hind)
mean(jääk)

#####

#tuleks logaritmida enne sobitamist
loghind=log(hind)
loghind
plot(loghind,col=2,xlab="Aeg",ylab="Soome Elektrienergia logaritmitud hind EUR/MWh")

log_lintrend=lm(loghind~aeg) #üks võimalik viise
#kasutame arima
log_lintrend=arima(loghind,xreg=aeg)
log_lintrend

#trendiks on : 3.0164+6e^-0.3*t;
log_jääk=residuals(log_lintrend)
#jääb järgi kui indeks maha võtta: jäägid=rida-trend => trend=rida-jäägid
log_trend=loghind-log_jääk
lines(log_trend,col="4")

```

```

plot(log_jääk, col=2,xlab="aeg",ylab="Elektrienergia hind eur/MWh pärast lineaarse trendi
eemaldamist")

korrelatsioonid(log_jääk,50)

PP.test(log_jääk)

# mudel

m4=arima(loghind,xreg=aeg,order=c(1,0,0))

tsdiag(m4,50)

#sobib väga hästi

m5=arima(loghind,xreg=aeg,order=c(1,0,1))

tsdiag(m5,50)

#sobib

m6=arima(loghind,xreg=aeg,order=c(1,0,1),seasonal=c(0,0,2))

tsdiag(m6,50)

#sobib

#parim m4

m4=arima(loghind,xreg=aeg,order=c(1,0,0),include.mean=F)

#####

##mõõdikud

mõõdikud=function(mudel,rida){

  vead=residuals(mudel) #mudel$x on esialgse aegrea andmed

  MAD=mean(abs(vead)) #keskmise absoluutviga, SASis MAE

  MSE=mean(vead**2) #keskmise ruutviga

  RMSE=sqrt(MSE)

  MAPE=mean(abs(vead/rida)) #keskmise suhteline viga, sageli korrutatakse 100ga, et saada
protsentides

  return(list(MAD=MAD,MSE=MSE,RMSE=RMSE,MAPE=MAPE))

}

mõõdikud(m0,hind)

###logaritmitud andmetel sobitatud mudeli mõõdikute leidmine

```



```

sqrt(mean(residuals(m4)**2))

prognoosid_log=loghind-residuals(m4) #prognoosid log jaoks

prognoosid=exp(prognoosid_log+m4$sigma2/2)

prognoosid

mudel=hind-prognoosid

MAD=mean(abs(mudel))

MSE=sqrt(mean((mudel)**2)) #leida eraldi ja siis täita tabel.

RMSE=sqrt(MSE)

MAPE=mean(abs((mudel)/hind))

#####

###ennustamine

#lin_trendi ennustus ja graafikule lisamine

ennustus=predict(m0,newxreg=((n+1):(n+9)),n.ahead=9)

ennustus

###ülemise ja alumise tõenäosus piiri leidmine 95% tõen

yleminepiir=ennustus$pred+1.96*ennustus$se

aluminepiir=ennustus$pred-1.96*ennustus$se

###80% ülemine ja alumine veapiir

yleminepiir1=ennustus$pred+1.28*ennustus$se

aluminepiir1=ennustus$pred-1.28*ennustus$se

#kasutades lisapaketti forecast

forecast(m0, newxreg=((n+1):(n+9)),h=9, level=c(80,95))

plot(forecast(hind),col=2)

grid(NA,20,lwd = 2) #joonte lisamine

#####

####ARIMAX mudeli leidmine

temp_sademed=read.table("f:/UT/loputoo/keskmine_temp_sademed.csv",sep=";",header=F,skip=1,dec
=","")

```

```

head(temp_sademed)

temp=ts(temp_sademed$V3,start=2000,frequency=12)

graafikud(temp,xlab="Aeg",ylab="Kuu keskmine temperatuur")

      #perioodilisus sees ja autokor kahanevad aeglaselt nulli

      #teeme funktsiooni, mis ei arvestaks neg

#jätame 2013 a väärtused välja

hind1=ts(rida, start=2000,end=c(2012,12),frequency=12)

length(hind1)

#####

n=length(hind1)

aeg=1:n

lin_trend=lm(hind1~aeg)

#kasutame arima

lin_trend=arima(hind1,xreg=aeg)

lin_trend

jääk2=residuals(lin_trend) #trendita


mu=mean(temp)

vaartus=(temp-mu)**2


graafikud(vaartus)

      #autokorr kanevad ning hinna jaoks ei ole oluline kui palju oli eelmine talv külmem kui see


temp_mudel=arima(vaartus,order=c(7,0,0))

tsdiag(temp_mudel,50) #sobib , parem

temp_mudel1=arima(vaartus,order=c(5,0,1))

tsdiag(temp_mudel1,50)

      #sobib

temp_mudel

```

```

#hinna aegrida ühepikkuseks temp aegreaga

model=arima(jääk2,order=c(7,0,0),fixed=temp_mudel$coef)
ccf(residuals(model),residuals(temp_mudel))

# 3 esimest väljas

z=window(hind1,start=c(2000,3))
xreg1=window(vaartus,start=c(2000,1),end=c(2012,10)) #2-nihet
xreg2=window(vaartus,start=c(2000,2),end=c(2012,11)) #1 nihe
xreg3=window(vaartus,start=c(2000,3),end=c(2012,12)) #0
xreg4=window(aeg,start=3,end=n) #aeg
lopp_mudel=arima(z,xreg=data.frame(xreg1,xreg2,xreg3,xreg4))
tulemus=residuals(lopp_mudel)#trendita
graafikud(tulemus)

lopp_mudel=arima(z,order=c(1,0,0),xreg=data.frame(xreg1,xreg2,xreg3,xreg4) )
tsdiag(lopp_mudel,50)

#sobis

#####

#ilma 2 nihketa
xreg2=window(vaartus,start=c(2000,2),end=c(2012,11)) #1 nihe
xreg3=window(vaartus,start=c(2000,3),end=c(2012,12)) #2
xreg4=window(aeg,start=3,end=n) #aeg
lopp_mudel2=arima(z,xreg=data.frame(xreg2,xreg3,xreg4))
tulemus1=residuals(lopp_mudel2)#trendita
graafikud(tulemus1)
lopp_mudel2=arima(z,order=c(1,0,0),xreg=data.frame(xreg2,xreg3,xreg4) )
tsdiag(lopp_mudel2,50)

#sobib

```

```
#####
```

```
#teine funktsioon
```

```
absvaartus=abs(temp-mu)
```

```
graafikud(absvaartus)
```

```
abstemp_mudel=arima(absvaartus,order=c(8,0,0))
```

```
tsdiag(abstemp_mudel,50) #sobib
```

```
abstemp_mudel1=arima(absvaartus,order=c(2,0,2))
```

```
tsdiag(abstemp_mudel1,50)#ei sobi kuidagi
```

```
abs_mudel=arima(jääk2,order=c(8,0,0),fixed=abstemp_mudel$coef)
```

```
ccf(residuals(abs_mudel),residuals(abstemp_mudel))
```

```
z=window(hind1,start=c(2000,4))
```

```
xreg1=window(vaartus,start=c(2000,1),end=c(2012,9)) #3-nihet
```

```
xreg2=window(vaartus,start=c(2000,2),end=c(2012,10)) #1 nihe
```

```
xreg3=window(vaartus,start=c(2000,4),end=c(2012,12)) #0 nihe
```

```
xreg4=window(aeg,start=4,end=n)
```

```
lopp_mudel3=arima(z,xreg=data.frame(xreg1,xreg2,xreg3,xreg4))
```

```
tulemus=residuals(lopp_mudel3)#trendita
```

```
graafikud(tulemus)
```

```
lopp_mudel3=arima(z,order=c(1,0,0),xreg=data.frame(xreg1,xreg2,xreg3,xreg4))
```

```
tsdiag(lopp_mudel3,50)
```

```
lopp_mudel3
```

```
####kui jätame nihked 0, 1 ära
```

```

xreg3=window(vaartus,start=c(2000,4),end=c(2012,12)) #nihe 0
xreg4=window(aeg,start=4,end=n)
lopp_mudel4=arima(z,xreg=data.frame(xreg3,xreg4))
lopp_mudel4=arima(z,order=c(1,0,0),xreg=data.frame(xreg2,xreg3,xreg4))
tsdiag(lopp_mudel4,50)

#####

#kolmas funktsioon

keeruline=function(temp){
  return((temp-12)*(temp>12)+(5-temp)*(temp<5))
}
keeruline(temp)
graafikud(keeruline(temp))

keertemp_mudel=arima(keeruline(temp),order=c(12,0,0))
tsdiag(keertemp_mudel,50) #sobib
keertemp_mudel1=arima(keeruline(temp),order=c(11,0,0))
tsdiag(keertemp_mudel1,50)#sobib ja parim
keertemp_mudel1 #parim mudel

keeruline_mudel=arima(jääk2,order=c(11,0,0),fixed=keertemp_mudel1$coef)
ccf(residuals(keeruline_mudel),residuals(keertemp_mudel1))

#esimene väljas -> nihe on 0 ja 1.

z=window(hind1,start=c(2000,2))
xreg=window(keeruline(temp),start=c(2000,1),end=c(2012,11)) #nihkega 1
xreg1=window(keeruline(temp),start=c(2000,2),end=c(2012,12))# Onihe
xreg2=window(aeg,start=2,end=n)

lopp_mudel5=arima(z,xreg=data.frame(xreg,xreg1,xreg2))

```

```
tulemus2=residuals(lopp_mudel2)#trendita
graafikud(tulemus2)
```

```
lopp_mudel6=arima(z,order=c(0,0,0),xreg=data.frame(xreg,xreg1,xreg2))
tsdiag(lopp_mudel1,50)# ei sobi
```

```
lopp_mudel7=arima(z,order=c(1,0,0),xreg=data.frame(xreg,xreg1,xreg2))
tsdiag(lopp_mudel1,50)#sobib
```

```
#####
```

```
###sademed
```

```
sademed=ts(temp_sademed$V4,start=2000,frequency=12)
graafikud(sademed)
```

```
#
```

```
sademed_mudel=arima(sademed,order=c(11,0,0))
tsdiag(sademed_mudel,50) #sobib
```

```
#mudeldamine
```

```
sademed_mudel=arima(hind1,order=c(11,0,0),fixed=sademed_mudel$coef)
ccf(residuals(sademed_mudel),residuals(sademed_mudel1))
```

```
# mitu ccf väärtust väljas
```

```
# ülekandefunktsiooni eeldused pole täidetud
```

```
####korrigeeritud
```

```
sademed1=ts(temp_sademed$V5,start=2000,frequency=12)
sademed2=ts(temp_sademed$V6,start=2000,frequency=12)
```

```
graafikud(sademed1)
```

```
sademed_mudel2=arima(sademed1,order=c(5,0,0))
```

```
tsdiag(sademed_mudel2,50) #ei sobi
```

```
sademed_mudel3=arima(sademed1,order=c(10,0,0))
```

```

tsdiag(sademed_mudel3,50) #sobib

#mudeldamine

sademed_mudel4=arima(jääk2,order=c(10,0,0),fixed=sademed_mudel3$coef)
ccf(residuals(sademed_mudel4),residuals(sademed_mudel3))

##kolmas ja neljas

z=window(z,start=c(2000,5))#esimese nelja jaoks ei ole regressoreid
xreg1=window(sademed1,start=c(2000,2),end=c(2012,9)) #nihkega 3
xreg2=window(sademed,startt=c(2000,1),end=c(2012,8))#nihkega 4
xreg3=window(aeg,start=1,end=n-4)

sademed_mudel4.1=arima(z,xreg=data.frame(xreg1,xreg2,xreg3))
sademed_mudel4.1=residuals(sademed_mudel4.1)
graafikud(sademed_mudel4.1)

lopp_mudel4.2=arima(z,order=c(1,0,0),xreg=data.frame(xreg1,xreg2,xreg3))
tsdiag(lopp_mudel4.2,50)

#####

lopp_mudel4.3=arima(z,order=c(1,0,0),xreg=xreg3)
tsdiag(lopp_mudel4.3,50)

#####

graafikud(sademed2)

sademed_mudel4=arima(sademed2,order=c(9,0,0))
tsdiag(sademed_mudel4,50) #ei sobi

sademed_mudel4=arima(sademed2,order=c(0,0,0),seasonal=c(1,0,0))
tsdiag(sademed_mudel4,50)

#mudeldamine

sademed_mudel5=arima(jääk2,order=c(0,0,0),seasonal=c(1,0,0),fixed=sademed_mudel4$coef)
ccf(residuals(sademed_mudel5),residuals(sademed_mudel4))# ei avalda mõju #mudeldamine

sademed_mudel5=arima(hind1,order=c(0,0,0),seasonal=c(1,0,0),fixed=sademed_mudel4$coef)
ccf(residuals(sademed_mudel5),residuals(sademed_mudel4))# ei avalda mõju

#####

#Arimax mudelite võrdlemine ja parimaga ennustamine

mõõdikud(lopp_mudel2,hind1)#parim

```

```

mõõdikud(lopp_mudel4,hind1)

mõõdikud(lopp_mudel7,hind1)

mõõdikud(lopp_mudel4.3,hind1)

xreg2=window(vaartus,start=c(2000,2),end=c(2012,11)) #1 nihe
xreg3=window(vaartus,start=c(2000,3),end=c(2012,12)) #0
xreg4=window(aeg,start=3,end=n) #aeg

lopp_mudel2=arima(z,xreg=data.frame(xreg2,xreg3,xreg4))

tulemus1=residuals(lopp_mudel2)#trendita

graafikud(tulemus1)

lopp_mudel2=arima(z,order=c(1,0,0),xreg=data.frame(xreg2,xreg3,xreg4) )

tsdiag(lopp_mudel2,50)

#####ennustamine 2 järgnevat väärtust

ennustus=predict(lopp_mudel2,1,newxreg=data.frame(xreg2=window(temp,start=c(2012,12),end=c(20
13,1)),xreg3=window(temp,start=c(2012,11),end=c(2012,12)),xreg4=window(aeg,end=n)))

ennustus

yleminepiir=ennustus$pred+1.96*ennustus$se
aluminepiir=ennustus$pred-1.96*ennustus$se

###80% ülemine ja alumine veapiir

yleminepiir1=ennustus$pred+1.28*ennustus$se
aluminepiir1=ennustus$pred-1.28*ennustus$se

yleminepiir
aluminepiir

yleminepiir1
aluminepiir1

```


Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina _____Anni Niidumaa_____

(*autori nimi*)

(sünnikuupäev: _____48802082735_____)

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose

Soome elektrienergia hinna modelleerimine ja lühiajaline ennustamine

_____,

(*lõputöö pealkiri*)

mille juhendaja on _____Raul Kangro_____,

(*juhendaja nimi*)

1.1.reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;

1.2.üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.

3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus/Tallinnas/Narvas/Pärnus/Viljandis, **22.05.2013**